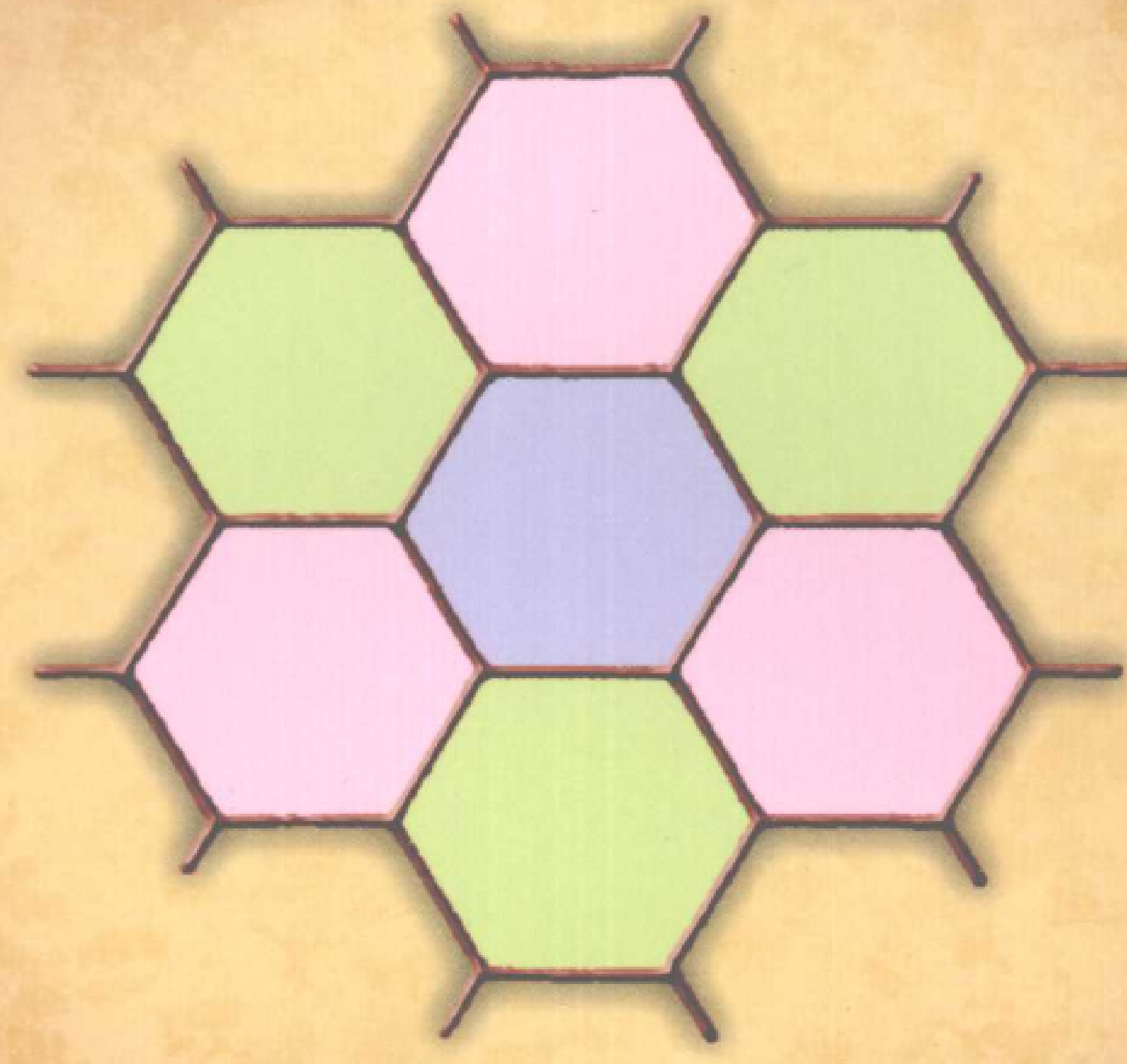
图灵数学·统计学丛书 24



A Survey of Modern Algebra

近世代数概论

(第 5 版)

[美] Garrett Birkhoff
Saunders Mac Lane 著

王连祥 徐广善 译

人民邮电出版社
POSTS & TELECOM PRESS

VISIT…

LANZAROTE
Caliente.com

图书在版编目(CIP)数据

近世代数概论：第 5 版 / (美) 伯克霍夫 (Birkhoff, G.), (美) 麦克莱恩 (Mac Lane, S.) 著；王连祥，徐广善译.
北京：人民邮电出版社，2008. 9
(图灵数学·统计学丛书)
书名原文：A Survey of Modern Algebra
ISBN 978-7-115-18387-3

I. 近… II. ①伯…②麦…③王…④徐… III. 抽象代数-高等学校-教材 IV. O153

中国版本图书馆 CIP 数据核字(2008)第 095136 号

内 容 提 要

本书出自近世代数领域的两位巨匠之手，是一本经典的教材. 全书共分为 15 章，内容包括：整数、有理数和域、多项式、实数、复数、群、向量与向量空间、矩阵代数、线性群、行列式与标准型、布尔代数与格、超限算术、环与理想、代数数域和伽罗瓦理论等.

本书适合数学专业及其他理工科专业高年级本科生和研究生使用，是一本非常有价值的教材和参考书.

图灵数学·统计学丛书

近世代数概论(第 5 版)

◆ 著　[美] Garrett Birkhoff　Saunders Mac Lane
译　王连祥　徐广善
责任编辑　明永玲

◆ 人民邮电出版社出版发行　北京市崇文区夕照寺街 14 号
邮编 100061　电子函件 315@ptpress.com.cn
网址：http://www.ptpress.com.cn
北京铭成印刷有限公司印刷

◆ 开本：700×1000　1/16
印张：27
字数：544 千字　2008 年 9 月第 1 版
印数：1–3 000 册　2008 年 9 月北京第 1 次印刷

著作权合同登记号　图字：01-2007-3613 号

ISBN 978-7-115-18387-3/O1

定价：69.00 元

读者服务热线：(010)88593802　印装质量热线：(010) 67129223

反盗版热线：(010)67171154

第 4 版前言

在本书第 1 版写完以来的 35 年间, 近世代数已成为全世界大学里的标准课程, 并且已有许多用于这门课程的教材. 尽管如此, 回顾一下我们在最初确定的基本指导思想 —— 也是现在这本书的基本指导思想 —— 看来是可取的.

“我们始终力求表达各种常用的定义的构思背景. 为此, 我们尽可能用较多的熟悉的例子说明每个新术语. 这在基础教材里特别重要, 因为它可以说明一切抽象概念都来源于对具体情况的分析.

“为了提高学生按照新概念独立思考的能力, 每个课题里我们都编入广泛多样的习题. 这些习题中, 一些用来计算, 一些用来进一步寻找新概念的例子, 另一些给出附加的理论推导. 后一种类型的习题对于学生熟悉正式证明的结构有重要的作用. 习题的选择使授课教师可根据情况取舍, 以适应大学本科生或一年级研究生不同程度的需要.

“近世代数也能够重新解释古典代数的结果, 使它们具有更大的统一性和一般性. 因此, 我们并不省略这些结果, 而努力把它们系统地编入近世代数的范围内.

“我们还力求不忽略如下事实: 对于许多学生来说, 代数学的意义在于它在其他领域的应用, 这些领域如高等分析、几何学、物理学和哲学等. 这使我们强调实数域和复数域、同抽象群相对照的变换群、对称矩阵及其对角化、正交群下和欧几里得群下的二次型分类, 并使我们最后加上布尔代数、格论和超限数的内容. 所有这些内容在数理逻辑和实函数近代理论中都很重要. ”

详细地说, 我们的第 1 章至第 3 章介绍交换环中线性方程和多项式方程理论, 在强调普通的整数环、有理数域的同时, 还强调了模 n 整数环和相伴多项式环. 第 4 章至第 5 章叙述实数域和复数域的基本代数性质, 这对于几何学和物理学具有头等重要性.

第 6 章通过群这个最简单最基本的概念, 引进非交换代数. 在第 7 章至第 10 章里, 群的概念系统地用到向量空间和矩阵上. 这里注意, 代数学在欧几里得几何、仿射几何和射影几何中一直起着最显著最基础的作用. 还讨论了对偶空间和张量积, 但不考虑推广到环上加法群.

第 11 章对布尔代数和格论的介绍完全重写了, 后面第 12 章, 是有关超限数的简短讨论. 最后的三章介绍了一般交换代数和算术: 理想和商环、域的扩张、代数数及其因子分解以及伽罗瓦理论.

许多章是相互独立的. 例如, 群论一章可以紧接第 1 章之后介绍, 而关于理想和

本书第 5 版是在 1977 年 Macmillan 出版的第 4 版基础上修订而成, 保留了第 4 版序, 未加新序.

—— 编者注

域的内容 (13.1 节和 14.1 节) 可以直接在向量空间后来研究.

这种独立性是为了使这本书既适用于只具备中学代数知识的学生的全年课程, 又适用于各式各样的短期课程. 例如介绍线性代数的一学期或小学期的课程, 可以以第 6 章至第 10 章为基础, 实数域和复数域是要强调的. 关于抽象代数的一学期课程, 可以安排第 1 章至第 3 章、第 6 章至第 8 章、第 11 章、第 13 章和第 14 章. 还可以有其他安排.

我们希望本书不仅继续作为教材, 而且为那些想要把近世代数的基本概念用于数学的其他分支 (包括统计学和计算), 用于物理学、化学和工程技术的读者作为方便的参考书.

在此愉快地向 C. 贝尔、A. A. 波恩涅、E. 阿廷、F. A. 菲肯、J. S. 弗雷姆、N. 雅各布森、W. 莱顿、G. 梅里曼、D. D. 米勒、I. 尼文以及许多其他朋友和同事致谢, 他们提供了有益的建议和改进. 另外还要感谢 S. 麦克莱恩夫人, 前三版中她做了许多事务性工作.

G. 伯克霍夫　于麻省剑桥
S. 麦克莱恩　于伊利诺伊芝加哥

目录

第 1 章 整数 ……1

1.1 交换环 · 整环 ……1
1.2 交换环的基本性质 ……2
1.3 有序整环的性质 ……7
1.4 良序原则 ……9
1.5 数学归纳法 · 指数定律 ……10
1.6 可除性 ……13
1.7 欧几里得算法 ……14
1.8 算术基本定理 ……18
1.9 同余式 ……20
1.10 环 $\mathbf{Z}_n$ ……23
1.11 集合 · 函数 · 关系 ……26
1.12 同构与自同构 ……29

第 2 章 有理数和域 ……31

2.1 域的定义 ……31
2.2 有理数域的构造 ……35
2.3 联立线性方程 ……39
2.4 有序域 ……43
*2.5 正整数公设 ……45
*2.6 皮亚诺公设 ……48

第 3 章 多项式 ……52

3.1 多项式形式 ……52
3.2 多项式函数 ……55
3.3 交换环的同态 ……59
*3.4 多元多项式 ……61
3.5 辗转相除法 ……63
3.6 单位与相伴 ……65
3.7 不可约多项式 ……67
3.8 唯一因子分解定理 ……69
*3.9 其他唯一因子分解整环 ……72
*3.10 爱森斯坦不可约判别准则 ……76
*3.11 部分分式 ……78

第 4 章 实数 ……82

4.1 毕达哥拉斯二难推论 ……82
4.2 上界与下界 ……83
4.3 实数公设 ……85
4.4 多项式方程的根 ……87
*4.5 戴德金分割 ……90

第 5 章 复数 ……94

5.1 复数的定义 ……94
5.2 复平面 ……96
5.3 代数基本定理 ……99
5.4 共轭数与实多项式 ……102
*5.5 二次方程与三次方程 ……104
*5.6 四次方程的根式解法 ……106
*5.7 稳定型方程 ……107

第 6 章 群 ……109

6.1 正方形的对称 ……109
6.2 变换群 ……111
6.3 其他例子 ……115
6.4 抽象群 ……117
6.5 同构 ……120
6.6 循环群 ……123
6.7 子群 ……126
6.8 拉格朗日定理 ……128
6.9 置换群 ……131
6.10 偶置换与奇置换 ……134
6.11 同态 ……136
6.12 自同构 · 共轭元素 ……138
*6.13 商群 ……141
*6.14 等价关系与同余关系 ……144

第 7 章 向量与向量空间 ……147

7.1 平面向量 ……147
7.2 推广 ……148
7.3 向量空间与子空间 ……150
7.4 线性无关与维数 ……153
7.5 矩阵与行等价 ……157

7.6 线性相关的检验 …… 159
7.7 向量方程·齐次方程 …… 163
7.8 基与坐标系 …… 167
7.9 内积 …… 172
7.10 欧几里得向量空间 …… 174
7.11 标准正交基 …… 177
7.12 商空间 …… 179
*7.13 线性函数与对偶空间 …… 181
第 8 章 矩阵代数 …… 186
8.1 线性变换与矩阵 …… 186
8.2 矩阵加法 …… 192
8.3 矩阵乘法 …… 193
8.4 对角矩阵·置换矩阵·三角形矩阵 …… 198
8.5 长方矩阵 …… 201
8.6 逆矩阵 …… 205
8.7 秩与零度 …… 210
8.8 初等矩阵 …… 212
8.9 等价与标准型 …… 216
*8.10 双线性函数与张量积 …… 218
*8.11 四元数 …… 222
第 9 章 线性群 …… 226
9.1 基的变换 …… 226
9.2 相似矩阵与特征向量 …… 228
9.3 全线性群与仿射群 …… 233
9.4 正交群与欧几里得群 …… 236
9.5 不变量与标准型 …… 240
9.6 线性型与双线性型 …… 242
9.7 二次型 …… 245
9.8 全线性群之下的二次型 …… 247
9.9 全线性群之下的实二次型 …… 250
9.10 正交群之下的二次型 …… 252
9.11 仿射群和欧几里得群之下的二次型 …… 256
*9.12 酉矩阵与埃尔米特矩阵 …… 260
*9.13 仿射几何 …… 263
*9.14 射影几何 …… 270
第 10 章 行列式与标准型 …… 275
10.1 行列式的定义和基本性质 …… 275
10.2 行列式的乘积 …… 279
10.3 作为体积的行列式 …… 282
10.4 特征多项式 …… 286
10.5 极小多项式 …… 290
10.6 凯莱–哈密顿定理 …… 294
10.7 不变子空间与可约性 …… 295
10.8 第一分解定理 …… 299
10.9 第二分解定理 …… 301
10.10 有理标准型与若当标准型 …… 304
第 11 章 布尔代数与格 …… 307
11.1 基本定义 …… 307
11.2 定律：同算术定律类比 …… 308
11.3 布尔代数 …… 310
11.4 其他基本定律的推导 …… 313
11.5 布尔多项式的标准型 …… 315
11.6 半序 …… 318
11.7 格 …… 320
11.8 集合表示 …… 323
第 12 章 超限算术 …… 327
12.1 数与集合 …… 327
12.2 可数集 …… 329
12.3 其他基数 …… 331
*12.4 基数的加法与乘法 …… 334
*12.5 取幂 …… 335
第 13 章 环与理想 …… 338
13.1 环 …… 338
13.2 同态 …… 341
13.3 商环 …… 345
*13.4 理想的代数 …… 347
13.5 多项式理想 …… 350
*13.6 线性代数中的理想 …… 353
13.7 环的特征 …… 355
13.8 域的特征 …… 357
第 14 章 代数数域 …… 359
14.1 代数扩张与超越扩张 …… 359
14.2 域上的代数元素 …… 361
14.3 根的添加 …… 363

14.4　次数与有限扩张 ……………… 365
14.5　多重代数扩张 ……………… 368
14.6　代数数 ……………… 371
14.7　高斯整数 ……………… 374
14.8　代数整数 ……………… 377
14.9　代数整数的和与积 ……………… 379
14.10　二次代数整数的因子分解 …… 381
第 15 章　伽罗瓦理论 ……………… 385
15.1　方程的根域 ……………… 385
15.2　唯一性定理 ……………… 387
15.3　有限域 ……………… 388
15.4　伽罗瓦群 ……………… 391
15.5　可分多项式与不可分多项式 … 395
15.6　伽罗瓦群的性质 ……………… 397
15.7　子群与子域 ……………… 399
15.8　三次不可约方程 ……………… 402
15.9　五次方程的不可解性 ……………… 406
参考文献 ……………… 410
数学符号表 ……………… 413
索引 ……………… 416

第 1 章　整　　数

1.1　交换环 · 整环

近世代数第一次揭示了数学系统的多变性和丰富性. 我们将构造并研究许多这样的系统, 但是它们中最基本的是最古老的数学系统 —— 由所有正整数(全体) 组成的系统. 与其有关的, 稍大一点的系统是由所有整数 $0, \pm 1, \pm 2, \pm 3, \cdots$ 组成的集合 $\mathbf{Z}$. 因为它与近世代数中的其他系统极为相似, 所以我们的讨论就从它开始.

整数具有许多有趣的代数性质. 在这一章里, 我们将假定一些像公设那样特别明显的性质, 并通过逻辑推理由它们导出许多别的性质.

我们首先假定加法和乘法的 8 个公设. 这些公设不仅对于整数成立, 而且对于许多其他数系都成立. 例如所有有理数 (分数)、所有实数 (无限小数) 和所有复数. 这些公设对于多项式和任意已知区间上的连续实函数也成立. 对于系统 R, 当这 8 个公设成立时, 我们称 R 为交换环.

定义　设 R 是由元素 $a, b, c, \cdots$ 组成的集合, 在 R 上定义了任意两个元素 a 与 b(不同或相同) 的和 $a+b$ 及积 ab. 如果下列公设 (i)~(viii) 成立, 那么 R 称为交换环:

(i) 封闭性. 若 a 与 b 在 R 中, 则和 $a+b$ 及积 ab 在 R 中.

(ii) 唯一性. 若 R 中 $a=a'$ 且 $b=b'$, 则

$$a+b=a'+b' \text{ 以及 } ab=a'b'.$$

(iii) 交换律. 对 R 中一切 a 与 b,

$$a+b=b+a, \qquad ab=ba.$$

(iv) 结合律. 对 R 中一切 a, b, c,

$$a+(b+c)=(a+b)+c, a(bc)=(ab)c.$$

(v) 分配律. 对 R 中一切 a, b, c,

$$a(b+c)=ab+ac.$$

(vi) 零. R 包含元素 0, 使得

$$a+0=a, \quad \text{对 } R \text{ 中一切 } a \text{ 成立}.$$

(vii) 单位元素. R 包含元素 $1 \neq 0$, 使得

$$a1 = a, \quad \text{对 } R \text{ 中一切 } a \text{ 成立}.$$

(viii) 加法逆元素. 对 R 中每个 a, 方程

$$a + x = 0 \quad \text{在 } R \text{ 中有解 } x.$$

所有整数的集合 $\mathbf{Z}$ 满足这些公设, 这是我们熟知的, 例如, 交换律和结合律是这么熟悉, 以致在平常应用时无须明确提及它们, 就把 $a+b+c$ 表示相等的数 $a+(b+c)$ 和 $(a+b)+c$. (vi) 中指出的 0 的性质是数零的特性; 类似地, (vii) 中指出的 1 的性质是数 1 的特性. 因为这两个公设形式上是类似的, 所以我们可以说, 0 和 1 分别是加法和乘法的"单位元素". (vii) 中的假定 $1 \neq 0$ 排除了平凡的情形 (否则, 交换环将是仅由整数 0 所组成的集合).

所有整数的系统 $\mathbf{Z}$ 具有另一个不能由上述公设推出的性质, 即若 $\mathbf{Z}$ 中 $c \neq 0$ 且 $ca = cb$, 则必有 $a = b$((ii) 中后一部分的逆性质). 但是交换环不一定都具有这个性质, 例如由已知区间上的全体实函数组成的集合, 虽然它们构成交换环, 但并不满足上述性质. 因此, 全体整数不仅构成交换环, 而且构成按下述意义定义的整环.

定义 满足下面附加公设的交换环是整环:

(ix) 消去律. 若 $c \neq 0$, 且 $ca = cb$, 则 $a = b$.

整环$\mathbf{Z}[\sqrt{2}]$. 由所有形为 $a + b\sqrt{2}$ 的数组成的整环是数论所感兴趣的, 这里 a 和 b 是普通整数 (在 $\mathbf{Z}$ 中). 在 $\mathbf{Z}[\sqrt{2}]$ 中,$a + b\sqrt{2} = c + d\sqrt{2}$ 当且仅当 $a = c, b = d$. 加法和乘法分别定义为

$$(a + b\sqrt{2}) + (c + d\sqrt{2}) = (a + c) + (b + d)\sqrt{2},$$

$$(a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (ad + bc)\sqrt{2}.$$

对于这些运算, 唯一性和交换律是容易验证的, 而 $0 + 0\sqrt{2}$ 相当于零, 并且 $1 + 0\sqrt{2}$ 相当于单位元素. $a + b\sqrt{2}$ 的加法逆元素是 $(-a) + (-b)\sqrt{2}$. 结合律和分配律的验证稍长一些, 消去律的验证将放到 1.2 节末尾.

1.2 交换环的基本性质

在初等代数中, 人们常常认为上述公设及其基本推论是理所当然的. 倘若对照特殊的例子检验代数运算时, 一般不会发生大的错误. 然而, 当我们想要得到对于整个代数系统都正确的结论时 (例如, 一般地, 对一切整环都成立), 必须多加小心. 我们必须确信, 所有证明只用到明显列出的公设和一般逻辑法则, 其中最基本的逻辑法则是相等关系的三个基本定律:

自反律 $a = a$.

对称律 若 $a=b$, 则 $b=a$.

传递律 若 $a=b$ 且 $b=c$, 则 $a=c$, 对一切 a, b 和 c 都成立.

现在我们列出几个在任意交换环 R 中都成立的法则, 并给出它们正式的证明.

法则 1 对 R 中一切 a,b,c, 有

$$(a+b)c=ac+bc.$$

这个法则可称为*右分配律*. 相对应地, 公设 (v) 是左分配律.

证明 对 R 中一切 a,b,c, 有

1° $(a+b)c=c(a+b)$ (乘法交换律),

2° $c(a+b)=ca+cb$ (分配律),

3° $(a+b)c=ca+cb$ ($1^\circ,2^\circ$, 传递律),

4° $ca=ac, cb=bc$ (乘法交换律),

5° $ca+cb=ac+bc$ (4°, 加法唯一性),

6° $(a+b)c=ac+bc$ ($3^\circ,5^\circ$, 传递律).

法则 2 对 R 中一切 $a, 0+a=a$, 且 $1\cdot a=a$.

证明 对 R 中一切 a, 有

1° $0+a=a+0$ (加法交换律),

2° $a+0=a$ (零的性质),

3° $0+a=a$ ($1^\circ,2^\circ$, 传递律).

$1\cdot a=a$ 的证明类似.

法则 3 如果 R 中的 z 具有性质 "对 R 中一切 $a, a+z=a$", 那么 $z=0$.

这个法则表明, R 仅包含一个 0 元素, 它可以起加法单位元素的作用.

证明 因为 $a+z=a$ 对一切 a 都成立, 所以当 a 为 0 时等式也成立.

1° $0+z=0$,

2° $0=0+z$ (1°, 对称律),

3° $0+z=z$ (法则 2, 当 a 为 z),

4° $0=z$ ($2^\circ,3^\circ$, 传递律).

在以后的这类证明中, 相等的对称律和传递律的反复运用, 我们都不必写出.

法则 4 对 R 中一切 a,b,c 成立:

$$\text{由 } a+b=a+c, \quad \text{可推出 } b=c.$$

这个法则称为加法消去律.

证明 根据公设 (viii), 对元素 a, 存在元素 x, 使 $a+x=0$. 因此

1° $x+a=a+x=0$ (加法交换律, 传递律),

2° $x=x, a+b=a+c$ (自反律, 假设),

3° $x+(a+b)=x+(a+c)$ (2°, 加法唯一性),

4° $b=0+b=(x+a)+b$

$=x+(a+b)=x+(a+c)$

$=(x+a)+c=0+c=c.$

(补上 4° 中每步的理由!)

法则 5 对每个 a, R 包含方程$a+x=0$ 的唯一解 x.

这个解通常用 $x=-a$ 表示. 因此这法则可被引述为 $a+(-a)=0$. 通常, 符号 $a-b$ 表示 $a+(-b)$.

证明 根据公设 (viii), 存在解 x. 如果 y 是第二个解, 那么根据传递律和对称律, $a+x=0=a+y$. 因此由法则 4, $x=y$. 证毕

法则 6 对 R 中给定的 a 和 b, 在 R 中存在唯一的 x, 使 $a+x=b$.

这个法则表明, 减法是可能的而且差是唯一的.

证明 取 $x=(-a)+b$. 则

$$a+x=a+[(-a)+b]=[a+(-a)]+b=0+b=b. \quad (\text{请给出理由!})$$

如果 y 是第二个解, 那么根据传递律 $a+x=b=a+y$, 因为由法则 4, $x=y$. 证毕

法则 7 对 R 中一切 a, $a\cdot 0=0=0\cdot a$.

证明

1° $a=a, a+0=a$ (自反律, 公设 (vi)).

2° $a(a+0)=aa$ (1°, 乘法唯一性).

3° $aa+a\cdot 0=a(a+0)=aa$ (分配律等).

$=aa+0$

4° $a\cdot 0=0$ (3°, 法则 4).

5° $0\cdot a=a\cdot 0=0$ (乘法交换律, 4°).

法则 8 如果 R 中的 u 具有性质 "对 R 中一切 a, $au=a$", 那么 $u=1$.

这个法则表明乘法单位元素 1 的唯一性. 证明类似于法则 3, 留作习题.

法则 9 对 R 中一切 a 和 b, $(-a)(-b)=ab$.

这个法则的特殊情形是 "玄" 律 $(-1)(-1)=1$.

证明 考察三重和(结合律!)

1° $[ab+a(-b)]+(-a)(-b)=ab+[a(-b)+(-a)(-b)]$.

由分配律、$-a$ 的定义、法则 7 和公设 (vi) 得

2° $ab+[a(-b)+(-a)(-b)]=ab+[a+(-a)](-b)$

$=ab+0(-b)=ab.$

同理, 有

3° $[ab+a(-b)]+(-a)(-b)=a[b+(-b)]+(-a)(-b)$

$$=a\cdot 0+(-a)(-b)=(-a)(-b).$$

因此, 根据相等的传递律和对称律, 从 1°, 2° 和 3° 得出结论. 证毕

其他各种简单而熟悉的法则, 都是我们公设的推论, 其中一些在下面习题中叙述.

另一个基本的代数定律在解二次方程时用到. 比如, 由 $(x+2)(x-3)=0$ 推出或者 $x+2=0$ 或者 $x-3=0$, 就用到这个定律, 它的一般形式就是断语:

$$\text{若}ab=0,\text{则或者}a=0\text{或者}b=0. \tag{1}$$

这个断语不是对一切交换环都成立的. 但是在任意整环 D 中, 根据消去律, 这个断语是正确的, 因为假设第一个因子不为零, 则 $ab=0=a\cdot 0$, 并且 a 可以消去, 因此 $b=0$. 反之, 在任意交换环 R 中, 从断语 (1) 可得到消去律, 因为如果 $a\neq 0, ab=ac$, 则有 $ab-ac=a(b-c)=0$, 由 (1) 得 $b-c=0$. 因此, 我们有

定理 1 *在交换环中, 乘法消去律等价于"非零因子之积不为零"这个命题.*

使乘积 $ab=0$ 的非零元素 a 和 b 有时称为"零因子", 因此, 交换环 R 中的消去律等价于"R 不包含零因子".

定理 1 可以用来证明 1.1 节末尾定义的整环 $\mathbf{Z}[\sqrt{2}]$ 的消去律, 如下所述. 假定 $\mathbf{Z}[\sqrt{2}]$ 包含零因子, 使

$$(a+b\sqrt{2})(c+d\sqrt{2})=(ac+2bd)+(ad+bc)\sqrt{2}=0.$$

由定义可推出 $ac+2bd=0, ad+bc=0$. 用 d 乘第一个等式, 用 c 乘第二个等式, 而后相减, 得到 $b(2d^2-c^2)=0$, 所以或者 $b=0$, 或者 $c^2=2d^2$. 如果 $b=0$, 则上述两个方程给出 $ac=ad=0$, 因此, 根据定理 1, 不是 $a=0$ 就是 $c=d=0$. 但是第一种情形 $a=0$ 意味着 $a+b\sqrt{2}=0$(因为 $b=0$); 第二种情形意味着 $c+d\sqrt{2}=0$, 所以这两种情形中, 都没有零因子.

现在余下 $c^2=2d^2$ 的情形, 这意味着 $\sqrt{2}=\dfrac{c}{d}$ 是有理数, 这是不可能的, 在 3.7 节定理 10 中将给出它的证明.

如果承认 $\sqrt{2}$ 是实数, 而且承认所有实数的集合构成整环, 那么借助于下面子整环的概念可以非常容易地证明 $\mathbf{Z}[\sqrt{2}]$ 是整环.

定义 *整环 D 的子整环是 D 的子集, 它对于同一种加法和乘法运算也是整环.*

显然, 子集 S 是子整环的充分必要条件是: S 包含 0 和 1; S 包含其中任意元素 a 的加法逆元素; S 包含其中任意两个元素 a 与 b 的和 $a+b$ 及积 ab.

习 题

对 $1\sim5$ 中的每个习题给出完整的证明，在证每一步时可用公设、前一步的结果、正文中已建立的法则或者已经作过的练习.

1. 证明下列法则在任意整环中都成立：
   (a) $(a+b)(c+d)=(ac+bc)+(ad+bd)$,
   (b) $a+[b+(c+d)]=(a+b)+(c+d)=[(a+b)+c]+d$,
   (c) $a+(b+c)=(c+a)+b$,
   (d) $a(bc)=c(ab)$,
   (e) $a[b+(c+d)]=(ab+ac)+ad$,
   (f) $a(b+c)d=(ab)d+a(cd)$.
2. (a) 证明法则 8. (b) 证明 $1\cdot1=1$.
   (c) 证明整环中仅有的幂等元素(即满足 $xx=x$ 的元素 x) 是 0 和 1.
3. 证明下列法则对任意整环中的 $-a$ 都成立：
   (a) $-(-a)=a$, (b) $-0=0$,
   (c) $-(a+b)=(-a)+(-b)$, (d) $-a=(-1)a$,
   (e) $(-a)b=a(-b)=-(ab)$.
4. 由习题 3(d) 和特殊情形 $(-1)(-1)=1$ 证明法则 9.
5. 证明在任意整环中下列法则对于运算 $a-b=a+(-b)$ 都成立：
   (a) $(a-b)+(c-d)=(a+c)-(b+d)$,
   (b) $(a-b)-(c-d)=(a+d)-(b+c)$,
   (c) $(a-b)(c-d)=(ac+bd)-(ad+bc)$,
   (d) $a-b=c-d$ 当且仅当 $a+d=b+c$,
   (e) $(a-b)c=ac-bc$.
6. 下列实数的集合是整环吗？为什么？
   (a) 所有偶数. (b) 所有奇数. (c) 所有正整数.
   (d) 所有实数 $a+b\sqrt[4]{5}$, 这里 a 和 b 为整数.
   (e) 所有实数 $a+b\sqrt[4]{9}$, 这里 a 和 b 为整数.
   (f) 所有分母为 2 的幂或 1 的有理数.
7. (a) 证明：仅由 0 和 1 组成的系统在通常的加法和乘法 (1+1=0(而不是 2) 除外) 运算之下是一个整环.
   (b) 证明：在仅由 0 组成的系统中定义 $0+0=0\cdot0=0$, 则除了 (vii) 中的条件 $0\neq1$ 外，它满足整环的所有公设.
8. (a) 证明：如果代数系统 S 满足整环的一切公设，(vii) 中的条件 $0\neq1$ 可能除外，那么，S 或者是整环，或者是仅由 0 组成的系统 (如习题 7(b) 中所描述的).
   (b) 在法则 1~9 的证明中用到条件 $0\neq1$ 吗？
9. 假定按通常定义任意两个整数的和，而任意两个整数的积定义为零. 在这两种运算之下，整环的公设中哪一些还仍然满足？

10. 找出两个函数 $f \neq 0$ 和 $g \neq 0$ 满足 $fg \equiv 0$.

1.3 有序整环的性质

因为所有普通整数的环 **Z** 在数学中起着独特的作用, 因此我们将研究它的特殊性质, 乘法交换律和消去律仅仅是其中两个. 许多其他性质都来源于整数有可能被排成通常的次序

$$\cdots -4, -3, -2, -1, 0, 1, 2, 3, 4, \cdots$$

这个次序常用关系 $a < b$ 来表达, 这里断语 "$a < b$" (a 小于 b) 意味着, 在上面所排的次序中, 整数 a 位于整数 b 的左边. 关系 $a < b$ 成立当且仅当差 $b-a$ 为正整数, 从而关系 $a < b$ 的每个性质可由正整数集合的性质导出. 因此我们假设正整数 $1, 2, 3, \cdots$ 的集合的下列三个性质作为公设.

加法律 两个正整数的和是正整数.

乘法律 两个正整数的积是正整数.

三分律 对于已知整数 a, 下面三种情况中有一个且仅有一个成立: 或者 a 为正整数, 或者 $a = 0$, 或者 $-a$ 为正整数.

顺便说一下, 在这些性质以及它们的推论中, 把 "正整数" 换成 "正有理数" 或 "正实数" 仍然成立. 为方便起见, 把包含具有这些性质的正元素的整环称为*有序整环*.

定义 *如果整环 D 中存在某些被称为正元素的元素, 它们满足类似于上面对整数指出的加法、乘法和三分律三个公设, 那么称 D 为有序整环.*

定理 2 *在任意有序整环中, 一切非零元素的平方都是正的.*

证明 设 a^2 已知, $a \neq 0$. 根据三分律, 或者 a 是正的, 或者 $-a$ 是正的. 在第一种情形中, 由正元素的乘法律知, a^2 是正的; 在第二种情形中, $-a$ 是正的, 因此根据 1.2 节的法则 9, $a^2 = (-a)^2 > 0$. 证毕

由此推出 $1 = 1^2$ 总是正的.

定义 *在有序整环中, $a < b$(读作 "a 小于 b") 和 $b > a$("b 大于 a") 这两个等价的说法都意味着 $b - a$ 是正的. 还有, $a \leqslant b$ 的意思是 $a < b$ 或者 $a = b$.*

根据这个定义, 正元素 a 现在可以描述为大于零的元素 a. 元素 $b < 0$ 称为*负元素*. 从上面的定义, 我们能推出关系 "小于" 的几个熟悉的性质.

传递律 若 $a < b$ 且 $b < c$, 则 $a < c$.

证明 根据定义, 由假设 $a < b$ 和 $b < c$ 可推出 $b - a$ 和 $c - b$ 是正的. 因此由加法律, 其和 $(b - a) + (c - b) = c - a$ 是正的, 这意味着 $a < c$.

正元素的三个基本公设对应着不等式的三个相应的性质.

不等式两边同时加上一元素 若 $a<b$, 则 $a+c<b+c$.

不等式两边同时乘以一正元素 若 $a<b$ 且 $c>0$, 则 $ac<bc$.

三分律 对任意 a 和 b, 三个关系式 $a<b, a=b$ 和 $a>b$ 中有一个且仅有一个成立.

作为例子, 我们证明第二个性质, 即一个不等式两边乘以正元素 c, 不等式仍然成立. 这结论要求我们证明 $bc-ac=(b-a)c$(参看 1.2 节的习题 5(e)) 是正的. 而这是乘法公设的直接推论, 因为根据假设, 因子 $b-a$ 和 c 都是正的. 类似地, 我们可以证明, 不等式两边乘以负元素时, 不等式反向 (参看下面的习题 1(c)).

定义 在有序整环中, 当元素 a 为 0 时, 它的绝对值 $|a|$ 是 0; 否则 $|a|$ 是元素对 $a, -a$ 中的正元素.

这个定义可以改述为

$$|a|=a, \quad \text{当} a \geqslant 0; \qquad |a|=-a, \quad \text{当} a<0. \tag{2}$$

适当地分这两种情形考虑, 我们可以证明和的绝对值与积的绝对值的定律:

$$|a+b| \leqslant |a|+|b|, \quad |ab|=|a||b|. \tag{3}$$

和的绝对值的定律也可以这样得到: 根据定义, 我们有

$$-|a| \leqslant a \leqslant |a| \quad \text{且} \quad -|b| \leqslant b \leqslant |b|,$$

因此, 把不等式相加可得

$$-(|a|+|b|) \leqslant a+b \leqslant |a|+|b|.$$

这立即表明, $a+b$ 不论是正的还是负的, 它的绝对值不能超过 $|a|+|b|$.

习 题

1. 从有序整环公设推导下列法则:
   (a) 若 $a<b$, 则 $a+c<b+c$, 反之亦真.
   (b) $a-x<a-y$ 当且仅当 $x>y$.
   (c) 若 $a<0$, 则 $ax>ay$ 当且仅当 $x<y$.
   (d) 若 $c>0$ 且 $ac<bc$, 则 $a<b$.
   (e) 若 $x+x+x+x=0$, 则 $x=0$
   (f) 若 $a<b$, 则 $a^3<b^3$.
   (g) 若 $c \geqslant 0$, 则由 $a \geqslant b$ 可推出 $ac \geqslant bc$.
2. 证明: 方程 $x^2+1=0$ 在有序数环中无解.

3. 尽你的可能，证明一些关于关系 $a \leqslant b$ 的定律.
4. 证明：在任意有序整环中，$||a|-|b|| \leqslant |a-b|$.

*5. [①]证明：在任意有序数环中，由 $a^7=b^7$ 可推出 $a=b$.

*6. 证明：在任意有序整环中，对一切 $a, b, a^2-ab+b^2 \geqslant 0$.

*7. 在整环 $\mathbf{Z}[\sqrt{2}]$ 中定义正元素，并证明加法、乘法和三分律三个公设成立.

*8. 设 D 为整环，在 D 中定义了关系 $a<b$，它满足正文中指出的传递律、不等式的加法和乘法原则以及三分律. 证明：当适当地选择正元素的集合时，D 为有序整环.

*9. 详细证明：有序整环的任一子整环为有序整环.

*10. 设 R 为任意交换环，它包含一个满足加法、乘法和三分律三个公设的正元素的子集. 证明 R 是有序整环.(提示：证明乘法消去律成立，分四种情况讨论：$x>0$ 且 $y>0$，$x>0$ 且 $-y>0$，$-x>0$ 且 $y>0$，$-x>0$ 且 $-y>0$.)

1.4 良序原则

如果有序整环 (如实数系那样) 的子集 S 的每个非空子集都包含最小元素，那么 S 称为*良序的*. 利用这个概念我们可以阐述整数的重要性质，该性质在特征上不是代数的，并且是其他数系所不具备的. 这就是

良序原则 全体正整数的集合是良序的.

换句话说，正整数的任意非空集合C 必包含某最小元素 m，使 C 中的 c 总有 $m \leqslant c$. 例如，最小正偶数是 2.

为了说明这个原则的作用，我们证明

定理 3 *0 和 1 之间没有整数.*

看一下全体整数的自然次序，这马上就清楚了. 但是我们想要指出，不看这个次序而从我们的假设出发也可以证明这个事实. 现在我们给出这个证明. 如果存在适合 $0<c<1$ 的任意整数 c，那么所有这种整数的集合 C 是非空的. 根据良序原则，这个集合中有最小整数 m，并且 $0<m<1$. 当我们用正数 m 乘这个不等式两边时，得到 $0<m^2<m$. 于是 m^2 是集合 C 中的另一整数，它小于已假定的 C 中最小元素 m. 这个矛盾导出定理 3 成立.

定理 4 *如果正整数的一个集合 S 包含 1，并且当它包含 n 时必包含 $n+1$，那么集合 S 包含任意正整数.*

证明 只须证明，由那些不含于 S 的正整数组成的集合 S' 是空的. 假设 S' 不是空的，它将包含最小元素 m. 但根据假设 $m \neq 1$，由此由定理 3，$m>1$，所以 $m-1$ 是正的. 但是 $1>0, m-1<m$，所以根据 m 的选择，$m-1$ 将在 S 中. 根据假设得到 $(m-1)+1=m$ 在 S 中. 这个矛盾使定理成立.

① 这里和后面较难的习题都打上了 * 号.

习 题

1. 证明：对任意整数 $a, a-1$ 是小于 a 的最大整数.
2. 下列集合中哪些是良序的：
   (a) 所有正奇数, (b) 所有负偶数,
   (c) 所有大于 -7 的整数, (d) 所有大于 249 的奇数.
3. 证明：良序集的任意子集是良序的.
4. 证明：如果整数的集合包含 -1000, 并且当它包含 x 时必包含 $x+1$, 那么这个集合包含所有正整数.
5. (a) 如果对整数集合 S 中的一切 x, 有整数 b, 使 $b \leqslant x$, 那么 S 称为有整数 b 作为"下界", b 本身不一定在 S 中. 证明：具有下界的任意非空整数集合 S 具有最小元素.
   (b) 证明：具有"上界"的任意非空整数集合具有最大元素.

1.5 数学归纳法 · 指数定律

现在我们可以按加法、乘法及序完整地列出全体整数集合的基本性质. 今后我们假定全体整数构成有序整环 $\mathbf{Z}$, 其中所有正元素的集合是良序的. 全体整数的集合的其他每个数学性质, 可以由此通过严格的逻辑推导来证明. 特别是, 我们能导出非常重要的

数学归纳法原理 设命题 $P(n)$ 与每个正整数 n 有关, 它或者正确或者错误. 如果 (i)$P(1)$ 是正确的, (ii) 对一切 k, 由 $P(k)$ 推出 $P(k+1)$, 那么 $P(n)$ 对一切正整数 n 都是正确的.

为了从良序假设导出这个原理, 只要观察使 $P(k)$ 正确的那些正整数 k 的集合, 因为它满足定理 4 的假设条件, 因此由定理 4 的结论便得到这个原则.

现在用归纳的方法来证明在任意交换环中成立的各种定律. 我们首先用它来形式地建立任意 n 个被加数的一般分配律

$$a(b_1+b_2+\cdots+b_n)=ab_1+ab_2+\cdots+ab_n. \tag{4}$$

为明确起见, 我们定义累加和 $b_1+\cdots+b_n$ 如下：

$$b_1+b_2+b_3=(b_1+b_2)+b_3,$$
$$b_1+b_2+b_3+b_4=[(b_1+b_2)+b_3]+b_4.$$

一般地, 可表为递推公式 (对于 $k \geqslant 1$)

$$b_1+\cdots+b_k+b_{k+1}=(b_1+\cdots+b_k)+b_{k+1}, \tag{5}$$

它表明, 如果对 k 项确定了括号的位置, 那么 $k+1$ 项中括号的位置由公式也可确定.

归纳证明 (4), 首先要证明 $n=1$ 时它正确, 这是显然的. 其次, 我们假定定律 (4) 对于 $n=k$ 正确, 要证明它对于 $n=k+1$ 正确. 根据定义 (5) 和分配律 (v),

$$\begin{aligned} a(b_1+\cdots+b_k+b_{k+1}) &= a[(b_1+\cdots+b_k)+b_{k+1}] \\ &= a(b_1+\cdots+b_k)+ab_{k+1}. \end{aligned}$$

右边第一项可以利用 (4) 对于 k 个被加数正确的假设化简, 于是上式化为

$$a(b_1+\cdots+b_k+b_{k+1}) = (ab_1+\cdots+ab_k)+ab_{k+1}.$$

因为根据定义 (5), 右边是 $ab_1+\cdots+ab_k+ab_{k+1}$, 所以我们完成了 (4) 的归纳证明.

类似的但更为复杂的归纳论证将得到一般结合律, 它断言: 和 $b_1+\cdots+b_k$ 或积 $b_1\cdots b_k$ 不管把括号括在哪里都有相同的值 (特殊情形出现在下面的习题 9). 应用这个结果和 (4), 我们还可建立双边一般分配律

$$\begin{aligned} &(a_1+\cdots+a_m)(b_1+\cdots+b_n) \\ &= a_1b_1+\cdots+a_1b_n+\cdots+a_mb_1+\cdots+a_mb_n. \end{aligned}$$

注意, 根据一般结合律和一般交换律, k 个已知项的和不管项的次序与分组如何总有相同的值.

任意交换环 R 中的正整指数也可以归纳处理. 如果 n 为正整数, 则幂 a^n 表示 n 个因子的积 $aa\cdots a$. 这也可叙述为递推定义

$$a^1=a, \quad a^{n+1}=a^na \qquad (\text{对}R\text{中任意}a), \tag{6}$$

根据这个公式, 就可以用已经算出的低次幂 a^n 来计算幂 a^{n+1}. 由这些定义, 我们可以对任意正整指数 m 和 n 证明下面常用的定律:

$$a^ma^n=a^{m+n}, \tag{7}$$

$$(a^m)^n=a^{mn}, \quad (ab)^m=a^mb^m. \tag{8}$$

例如, 第一个定律可以对 n 用归纳法证明. 当 $n=1$ 时, (7) 式变成 $a^ma=a^{m+1}$, 这正是 a^{m+1} 的定义. 其次假定定律 (7) 对于任何 m 和已知正整数 $n=k$ 是正确的, 并且考虑比 k 大 1 的指数 $k+1$ 的类似表达式 a^ma^{k+1}, 我们逐次应用定义、结合律、归纳假设和定义, 得到

$$a^ma^{k+1}=a^m(a^ka)=(a^ma^k)a=a^{m+k}a=a^{(m+k)+1}=a^{m+(k+1)},$$

这就是定律 (7) 的 $n = k + 1$ 的情形, 因此完成了归纳证明.

最后, 我们证明二项公式在任意交换环 R 上成立. 首先用递推公式

$$0! = 1 \quad 和 \quad (n+1)! = n!(n+1)$$

定义非负整数上的阶乘函数 $n!$. 然后对 $\mathbf{Z}$ 中的 $n \geqslant 0$, 类似地用

$$\binom{n}{0} = \binom{n}{n} = 1 \quad 和 \quad \binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}$$

定义二项系数. 由这些定义, 再对 n 用归纳法, 得到

$$\begin{aligned}(x+y)^n &= x^n + nx^{n-1}y + \cdots + \binom{n}{k}x^{n-k}y^k + \cdots + y^n \\ &= \sum_{k=0}^{n}\binom{n}{k}x^{n-k}y^k\end{aligned} \tag{9}$$

和

$$k!(n-k)!\binom{n}{k} = n!. \tag{10}$$

$\left(即\ \binom{n}{k} = \dfrac{n!}{k!(n-k)!}\right.$, 证明留作习题. $\left.\right)$

数学归纳法原理允许我们在证明 $P(n+1)$ 时, 随意假定 $P(n)$ 的正确性. 我们现在指出, 人们甚至可以对一切 $k \leqslant n$ 假定 $P(k)$ 的正确性. 这称为

数学归纳法第二原理 设命题 $P(n)$ 与每个正整数 n 有关, 如果对每个 m, 由假设 "$P(k)$ 对一切 $k < m$ 是正确的", 可以推出结论 "$P(m)$ 本身是正确的", 那么 $P(n)$ 对一切 n 都是正确的.

证明 设 S 是使 $P(n)$ 错误的正整数集合. 如果 S 不空, 则它有最小的数 m. 根据 m 的选法, $P(k)$ 对一切 $k < m$ 是正确的, 在此根据假设, $P(m)$ 本身必是正确的, 这就得出矛盾. 于是 S 只能是空的. 证毕

注意, 在 $m = 1$ 的情形中, 所有 $k < 1$ 的集合是空的, 因此必须暗含 $P(1)$ 的证明.

习 题

1. 用归纳法证明下列正指数定律在任意整环中成立:
   (a) $(a^m)^n = a^{mn}$, (b) $(ab)^n = a^n b^n$, (c) $1^n = 1$.
2. 用归纳法证明 $1 + 2 + \cdots + n = \dfrac{n(n+1)}{2}$.
3. 证明公式 (9) 和公式 (10).
4. 用归纳法证明: $x_1^2 + \cdots + x_n^2 > 0$, 除非 $x_1 = \cdots = x_n = 0$.

5. 用归纳法证明下列加法公式：

(a) $1+4+9+\cdots+n^2=\dfrac{n(n+1)(2n+1)}{6}$,

(b) $1+8+27+\cdots+n^3=\left[\dfrac{n(n+1)}{2}\right]^2$.

6. 证明：在任意有序整环中，负元素的任意奇次幂都是负的.
7. 用归纳法而不用良序原则证明定理 3.（提示：设 $P(n)$ 表示 $n\geqslant 1$.）
*8. 利用习题 7，由数学归纳法原理证明良序原则.（提示：设 $P(n)$ 为命题：任意包含一个 $\leqslant n$ 的数的一组正整数具有最小元素.）
9. 用定义 (5) 证明下面结合律：

$$(a_1+\cdots+a_m)+(b_1+\cdots+b_n)=a_1+\cdots+a_m+b_1+\cdots+b_n.$$

10. 给出两个函数之积的 n 阶导数公式，并对 n 用归纳法证明公式.
*11. 证明：对任何底 $a>1$，每个正整数 m 具有形如

$$a^nr_n+a^{n-1}r_{n-1}+\cdots+a^2r_2+ar_1+r_0$$

的唯一表达式，其中整数 r_k 满足 $0\leqslant r_k<a, r_n\neq 0$.
*12. 以下例说明习题 11：取 7 为底，变换方程 63×111=6993，并乘出来检验.
13. 药剂师仅有 1, 3, 9, 27 和 81 盎司五个砝码及双盘天平（砝码可放入任一盘中）. 证明他能够称出 1 ~ 121 盎司的任意重量.
14. 证明：任意 9 的倍数其各位数字之和可被 9 整除.

1.6 可 除 性

整系数方程 $ax=b$ 不总是有整数解 x. 如果有整数解，则称 b 可被 a 整除. 数论首先要研究的就是这个问题.

在任意整环中也有类似的可除性概念，定义如下：

定义 *在整环 D 中，如果有 D 中某一 q，使 $b=aq$，则称元素 b 可被元素 a 整除. 当 b 可被 a 整除时，我们记作 $a|b$. 我们又称 a 是 b 的因子，b 是 a 的倍数. 1 的因子称为 D 的单位或可逆元素.*

同相等关系 $a=b$ 一样，关系 $a|b$ 满足自反律和传递律：

$$a|a;\qquad 由\ a|b\ 和\ b|c\ 可推出\ a|c. \tag{11}$$

(11) 的第一个定律是显然的，因为 $a=a\cdot 1$ 意味着 $a|a$. 为证明第二个定律，回想一下整除的定义，$a|b$ 和 $b|c$ 意味着有某整数 d_1 和 d_2，使 $b=ad_1$ 和 $c=bd_2$，将第一个方程代入第二个方程中得出 $c=a(d_1d_2)$. 因为 d_1d_2 是整数，按照定义，这表明 $a|c$，同 (11) 中所断言的一样.

定理 5　$\mathbf{Z}$ 中仅有的单位是 ± 1.

这个定理实际上断言：对于整数 a 和 b, $ab = 1$ 意味着 $a = \pm 1$ 和 $b = \pm 1$. 根据积的绝对值定律, 由 $ab = 1$ 得出 $|ab| = |a| \cdot |b| = 1$. 因为 a, b 都不为零, 所以 $|a|$ 和 $|b|$ 是正数. 由于 0 与 1 之间没有正整数 (定理 3), 因此根据三分律, $|a| \geqslant 1$ 和 $|b| \geqslant 1$. 这两个不等式随便哪个不等关系成立, 积 $|a||b|$ 就不可能是 1. 因此 $|a| = |b| = 1$, 即如定理所言, $a = \pm 1, b = \pm 1$.

推论　如果整数 a 和 b 彼此可整除 ($b|a$ 且 $a|b$), 那么 $a = \pm b$.

证明　根据假设 $a = bd_1$ 且 $b = ad_2$, 因此 $a = ad_2d_1$. 如果 $a = 0$, 则 $b = 0$, 结论当然成立. 如果 $a \neq 0$, 消去 a 后得到 $1 = d_2d_1$. 那么, 根据定理 5, $d_1 = \pm 1$, 因此 $a = \pm b$.　　证毕

因为 $a = a \cdot 1 = (-a)(-1)$, 所以任意整数 a 可被 $a, -a, 1$, 和 -1 整除.

定义　如果整数 p 不为 0 或 ± 1, 并且 p 只能被 ± 1 和 $\pm p$ 整除, 那么称 p 为素数.

前几个正素数是

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31.$$

不是 1 或素数的任何正整数都可分解成素因子的积, 例如

$$128 = 2^7, \qquad 90 = 9 \times 10 = 3^2 \times 2 \times 5,$$
$$672 = 7 \times 96 = 7 \times 12 \times 8 = 7 \times 3 \times 2^5.$$

经验表明, 不论怎样进行分解, 总会得到相同的素因子. 这种素因子分解的唯一性可以用下面我们将讨论的最大公因子来证明.

习　　题

1. 证明任意整环 D 中单位的下列各性质：
   (a) 两个单位之积是单位.
   (b) D 的单位 u 可整除 D 中每个元素.
   (c) 若 c 整除 D 中每个 x, 则 c 是单位.
2. 证明：若 $a|b$ 且 $a|c$, 则 $a|(b+c)$.
3. 证明：若 $b > 1$, 而且不是素数, 则它有正的素因子 $d \leqslant \sqrt{b}$.
4. 列出所有小于 100 的正素数. (提示：删去 2, 3, 5, 7 的所有倍数, 并应用习题 3.)
5. 设 $a|b$, 证明：当 $b \neq 0$ 时, $|a| \leqslant |b|$.

1.7　欧几里得算法

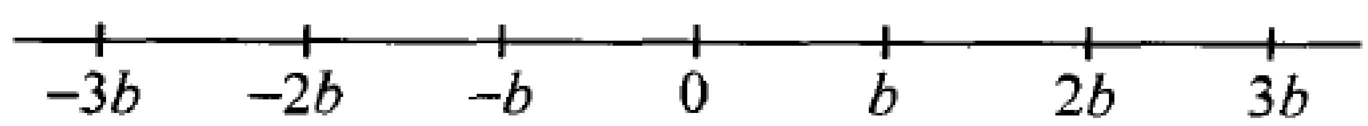
整数 a 除以 b 用普通的除法就得到商 q 和余数 r. 正式地说, 这相当于下面的

断语.

除法算式 对于给定的整数 a 和 $b, b>0$, 存在整数 q 和 r, 使得

$$a=bq+r, \quad 0 \leqslant r<b. \tag{12}$$

几何描述 如果我们设想全部数都显示在实轴上, 那么 b 的所有可能倍数 bq 构成直线上等距分点的集合, 对应于 a 的点必落在由这些点确定的区间中的一个, 比如说, 落在 bq 和 $b(q+1)$ 之间的区间中, 右端点除去. 这意味着 $a-bq=r$, 其中 r 表示短于区间全长 b 的长度, 因此, 如断言所述 $0 \leqslant r<b$, 这个描述启发我们用良序的性质进行以下证明.



证明 当然存在 b 的某整数倍不超过 a, 例如, 因为 $b>0$, 根据定理 3, $b \geqslant 1$, 所以 $(-|a|)b \leqslant -|a| \leqslant a$. 因此差 $a-bx$ 的集合至少包含一个非负整数, 即 $a-(-|a|)b$. 从而, 根据良序的性质, 存在最小非负的 $a-bx$, 比如说, $a-bq=r$. 由构造可知 $r \geqslant 0$. 而当 $r \geqslant b$ 时, 则 $a-b(q+1)=r-b \geqslant 0$ 将小于 $a-bq$, 这与我们对 q 的选择相违背. 由此得出结论: 当 $a=bq+(a-bq)=bq+r$ 时, $0 \leqslant r<b$.

推论 1 对给定的整数 a 和 b, 满足 (12) 的商 q 和余数r 是唯一确定的.

证明 假设 $a=bq+r=bq'+r'$, 式中 $0 \leqslant r<b, 0 \leqslant r'<b$, 那么 $r-r'=b(q'-q)$ 在数值上小于 b, 但它是 b 的倍数. 因此 $r-r'$ 必为零, 所以 $r=r', bq=bq', q=q'$, 这就得出 q 和 r 的唯一性. 证毕

我们常常有必要不涉及单个的整数而去处理某整数集合, 像由 3 的所有倍数组成的集合:

$$\cdots,-6,-3,0,3,6,9,\cdots,$$

这个集合具有重要性质: 集合中的任意两个整数的和或差仍然是集合中的整数. 一般地, 如果整数集合 S 包含 S 中任意两个整数 a 与 b 的和 $a+b$ 及差 $a-b$, 则称集合 S 在加法与减法之下是封闭的. 所有偶数 (正的、负的和零) 构成这样的集合. 更一般地, 任意固定的整数 m 的所有倍数 xm 的集合在加法与减法之下是封闭的, 这因为 $xm \pm ym=(x \pm y)m$ 是 m 的倍数. 我们现在证明: 这种倍数的集合是具有这些性质的唯一的整数集合.

定理 6 在加法与减法之下封闭的任意非空整数集合, 不是仅由零组成, 就是包含最小正整数并由这个整数的所有倍数组成.

证明 设这样的集合 S 包含元素 $a \neq 0$. 则 S 包含差 $a-a=0$, 因此包含差 $0-a=-a$. 所以 S 中至少有一个正元素 $|a|= \pm a$. 根据良序原则, S 中存在最小正元素 b.

集合 S 必包含 b 的所有整倍数. 这是因为我们首先可用归纳法 (对 n) 证明 b 的任何正倍数 nb 在 S 中: 若 $n=1$, 则 b 在 S 中; 若已知 kb 在 S 中, 则 $(k+1)b=kb+b$ 是 S 的两个元素之和, 因此在 S 中. 而 b 的任何负倍数 $(-n)b=0-(nb)$ 是 S 的两个元素之差, 因此也在 S 中.

集合 S 只能包含 b 的所有整倍数. 这是因为如果 a 是 S 的任意元素, 则由除法算式可得出差 $a-bq=r$, 它也在 S 中. 余数 r 非负且小于 b, 而 b 是 S 中的最小正元素, 因此 $r=0, a=bq$ 是 b 的倍数, 如断言所述. 证毕

定义 如果整数 d 是整数 a 和 b 的公因子, 并且是任何其他公因子的倍数, 那么称 d 为 a 和 b 的最大公因子 (g.c.d.). 用符号表示, d 必有性质

$$d|a; \qquad d|b; \qquad \text{由} \quad c|a \text{ 和 } c|b \quad \text{可推出} \quad c|d.$$

例如 3 和 −3 都是 6 和 9 的最大公因子. 按照定义, 两个不同的最大公因子必彼此整除, 因此它们仅相差一个符号. a 和 b 的两个可能的最大公因子 $\pm d$ 中, 正的最大公因子常用符号 (a,b) 表示. 值得注意的是, 最大公因子定义中的形容词 "最大", 主要不是指 d 的数值比任何其他公因子 c 大, 而是指 d 为任何这种 c 的倍数.

定理 7 任意两个整数 $a\neq 0$ 和 $b\neq 0$ 有正的最大公因子 (a,b). 它可表为 a 和 b 的具有整系数 s 和 t 的线性组合, 形为

$$(a,b)=sa+tb. \tag{13}$$

证明 考虑形为 $sa+td$ 的数. 对于任意两个这样的数:

$$(s_1a+t_1b)\pm(s_2a+t_2b)=(s_1\pm s_2)a+(t_1\pm t_2)b.$$

所有, 所有整数 $sa+tb$ 的集合 S 在加法和减法之下是封闭的. 因此根据定理 6, S 由某一最小正整数 $d=sa+td$ 的所有倍数组成. 由此公式显然可知, a 和 b 的任何公因子必是 d 的因子. 另一方面, 原来的整数 $a=1\cdot a+0\cdot b$ 和 $b=0\cdot a+1\cdot b$ 都在所考虑的集合 S 之中, 因此必为这个集合最小整数 d 的倍数, 换句话说, d 是公因子. 因此它就是所要求的最大公因子. 证毕

类似地, a 和 b 的公倍数的集合 M 在加法和减法之下是封闭的. 它的最小正元素 m 将是 a 和 b 的公倍数, 它整除每个公倍数. 于是 m 是最小公倍数(或 l.c.m.).

定理 8 任意两个整数 a 和 b 有最小公倍数 $m=[a,b]$, 它是 a 和 b 的每个公倍数的因子, 并且它自己是 a 和 b 的公倍数.

为找到两个整数 a 和 b 的最大公因子的明显表达式, 可应用所谓欧几里得 (Euclid) 算法. 我们可以假定 a 和 b 都是正的, 因为负整数 b 可用 $-b$ 代替, 并不改变最大公因子 $(a,b)=(a,-b)$. 除法算式给出

$$a=bq_1+r_1, \qquad 0\leqslant r_1<b. \tag{14}$$

整除 a 和 b 的每个整数必整除余数 r_1; 反之, (14) 中 b 和 r_1 的每个公因子是 a 的因子, 所以 a 和 b 的公因子同 b 和 r_1 的公因子一样, 因此最大公因子 (a,b) 和 (b,r_1) 相等. 这种简化可以在 b 和 r_1 的位置上反复进行:

$$\begin{aligned}
b &= r_1q_2 + r_2, & 0 &< r_2 < r_1,\\
r_1 &= r_2q_3 + r_3, & 0 &< r_3 < r_2,\\
&\vdots & &\vdots\\
r_{n-2} &= r_{n-1}q_n + r_n, & 0 &< r_n < r_{n-1},\\
r_{n-1} &= r_nq_{n+1}.
\end{aligned} \tag{15}$$

因为余数不断减小, 最后必有余数 r_{n+1} 为零[①], 正像我们在最后一个方程中表示的那样. 以上论证表明, 所要求的最大公因子是

$$(a,b) = (b,r_1) = (r_1,r_2) = \cdots = (r_{n-1},r_n).$$

但是 (15) 的最后一个方程表明 r_n 本身是 r_{n-1} 的因子, 因此最后一个最大公因子恰是 r_n 自己. 于是已知整数 a 和 b 的最大公因子是欧几里得算法 (14) 和 (15) 中最后一个非零余数 r_n.

利用欧几里得算法, 也可把最大公因子明显地表示为线性组合 $sa+tb$. 这只要用 a 和 b 表示逐次的余数 r_i 就可做到. 例如

$$\begin{aligned}
r_1 &= a - bq_1 = a + (-q_1)b,\\
r_2 &= b - q_2r_1 = (-q_2)a + (1+q_1q_2)b,\\
&\vdots \qquad\qquad\quad \vdots
\end{aligned}$$

这些方程的形式表明, 我们最后能把 r_n 表为 a 和 b 的线性组合, 它具有包含商 q_i 的整系数 s 和 t.

最大公因子的表达式 $(a,b)=sa+tb$ 非常有用. 一个重要的推论是, 整除两个数之积的素数必至少整除其中一个因子:

定理 9 *如果 p 为素数, 那么由 $p|ab$ 可推出 $p|a$ 或 $p|b$.*

证明 根据素数定义, p 只有因子 ±1 和 $\pm p$. 如果结论 $p|a$ 是错误的, 则 p 和 a 的公因子只能是 ±1, 因此 1 是 a 和 p 的最大公因子, 并可表为 $1=sa+tp$. 上式两边都乘以 b, 我们有

$$b = sab + tbp.$$

右边两项可被 p 整除, 因此左边 b 可被 p 整除, 这就是定理中的第二种可能性.

证毕

如果 $(a,b)=1$, 那么我们称 a 和 b *互素*. 换句话说, 如果两个整数 a 和 b 没有 ±1 以外的公因子, 则它们互素. 用来证明定理 9 的方法也可证明下面的推广:

---

① 为什么? 其证明包含良序原则吗?

定理 10 如果 $(c, a) = 1$ 且 $c|ab$, 那么 $c|b$.

对于两个互素的整数 a 和 c, 如果整数 m 是它们每一个的倍数, 则我们可推出下面的一个结果. 因为这样的 m 有形式 $m = ad$, 并可被 c 整除, 所以根据定理 10, 有 $c|d, m = ad = a(cd')$, 因此乘积 ac 可整除 m. 这就证明了

定理 11 如果 $(a, c) = 1, a|m$ 且 $c|m$, 那么 $ac|m$.

习 题

1. 利用欧几里得算法求最大公因子：
   (a) (14, 35), (b) (11, 15), (c) (180, 252),
   (d) (2873, 6643), (e) (4148, 7684), (f) (1001, 7655).
2. 把习题 1(a), (b), (c) 中的 (x, y) 写成形式 $sx + ty$(s, t 为整数).
3. 证明：对任意整数 $a, (0, a) = |a|$.
4. 如果 $a > 0$, 证明 $(ab, ac) = a(b, c)$.
5. 证明：由 $b|c$ 和 $|c| < b$ 推出 $c = 0$. (这个事实已用于证明推论 1.)
6. (a) 证明：任意三个整数 a, b, c 有最大公因子, 它可表成

$$sa + tb + uc.$$

   (b) 证明 $((a, b), c) = (a, (b, c)) = ((a, c), b)$.
7. 讨论习题 3 ~ 习题 5 和习题 6(b) 关于最小公倍数的情形.
8. 证明：在减法之下封闭的整数集合, 在加法之下也必是封闭的.
9. 证明：仅在加法之下封闭的整数集合不一定由一个固定元素的所有倍数组成.
10. 在欧几里得算法中, 对 k 用归纳法证明：每个余数可表成 $r_k = s_k a + t_k b$, 式中 s_k 和 t_k 为整数.
11. 给出定理 10 的详细证明.

*12. 证明：对任意正整数 a, b, 所有 $ma + nb$(m, n 为正整数) 的集合包含大于 ab 的 (a, b) 的所有倍数.

13. 如果 q 为整数, 使得对一切整数 a 和 b, 由 $q|ab$ 可推出 $q|a$ 或 $q|b$, 证明：q 是 $0, \pm 1$ 或素数 (参考定理 9).
14. (a) 证明：若 $(a, m) = (b, m) = 1$, 则 $(ab, m) = 1$.
    (b) 证明：若 $(a, c) = d, a|b$ 且 $c|b$, 则 $ac|bd$.
    (c) 证明 $[a, c] = \frac{ac}{(a,c)}$.

1.8 算术基本定理

现在我们容易证明整数唯一因子分解定理, 它也称为算术基本定理.

定理 12 任意非零整数可表为单位 (± 1) 乘以正素数的积. 如果不计素因子出现的顺序, 这种表示是唯一的.

证明 任意整数 a 能写成这样的乘积, 可以用逐次把 a 分解成较小因子的办法来证明. 证明中用到数学归纳法第二原理, 描述如下. 显然只考虑正整数 a 就够了.

设 $P(a)$ 为命题: a 能像定理 12 中所说的那样分解. 如果 $a=1$ 或 a 为素数, 则 $P(a)$ 当然正确. 另一方面, 如果 a 是复合数, 那么 a 有一个正因子 b, 它不是 1 也不是 a, 因此 $a=bc$, 其中 $b<a, c<a$. 但是根据数学归纳法第二原理, 我们可假定 $P(b)$ 和 $P(c)$ 正确, 所以 b 和 c 可表为素数之积:

$$b=p_1p_2\cdots p_r,\qquad c=q_1q_2\cdots q_s,$$

对于 a, 由上式得出复合数的表达式

$$a=bc=p_1p_2\cdots p_rq_1q_2\cdots q_s,$$

这就是所要求的形式.

为证明唯一性, 我们必须考虑整数 a 的两个可能的素因子分解

$$a=(\pm 1)p_1p_2\cdots p_m=(\pm 1)q_1q_2\cdots q_n.$$

因为素数 p_i 和 q_j 都是正的, 所以两种分解中的项 ± 1 必须一致. 第一个因子分解式中的素数 p_1 是乘积 $a=\pm q_1q_2\cdots q_n$ 的因子. 因此反复应用定理 9, 就有 p_1 必至少整除这个乘积的一个因子 q_j. 因为 $p_1|q_j$, 并且二者都是正素数, 所以 $p_1=q_j$. 重新排列分解式 $q_1q_2\cdots q_n$, 使 q_j 第一个出现, 那么 p_1 与 q_j 相消, 留下

$$p_2p_3\cdots p_m=q_2'q_3'\cdots q_n',$$

式中符号 “′” 表示这些 q 的新的顺序. 继续这个过程直到所得方程的一边没有留下素因子. 此时方程的另一边也没有素因子. 所以在原来的因子分解式中, $m=n$. 把第二个因子分解式中的素因子重新排列, 我们就使两个因子分解式完全一致, 这同唯一性定理中所断言的一样. 证毕

数的因子分解中, 同一个素数 p 可以出现几次. 把所出现的相同素数集中起来, 分解式可写为:

$$a=\pm p_1^{e_1}p_2^{e_2}\cdots p_k^{e_k}\qquad (1<p_1<p_2<\cdots<p_k). \tag{16}$$

由唯一性定理可知: 每个素数 p_i 的指数 e_i 由给定的数 a 唯一确定.

习 题

1. 描述求两个整数的最大公因子和最小公倍数的系统过程, 这两个整数的素数幂分解式 (16) 是已知的. 以 $a=216, b=360$ 和 $a=144, b=625$ 为例加以说明. (提示: 对于能整除 a 或 b 中的一个而不能整除两个的素数, 采用 “哑” 零分量是有益的.)

2. 如果 $V_p(a)$ 表示能整除非零整数 a 的素数 p 的最高次幂的指数, 证明公式
   (i) $V_p(a+b) \geqslant \min\{V_p(a), V_p(b)\}$, (ii) $V_p((a,b)) = \min\{V_p(a), V_p(b)\}$,
   (iii) $V_p(ab) = V_p(a) + V_p(b)$, (vi) $V_p([a,b]) = \max\{V_p(a), V_p(b)\}$.
3. 如果 $||a|| = 2^{-V_p(a)}$, 其中 $V_p(a)$ 的涵义同习题 2. 证明

$$||ab|| = ||a|| \cdot ||b|| \text{和} ||a+b|| \leqslant \max\{||a||, ||b||\}.$$

*4. 设 $V(a)$ 是对一切非零整数 a 有定义的非负整值函数, 并且具有习题 2 中的性质 (i) 和 (iii). 证明: $V(a)$ 或者恒为零, 或者是习题 2 中的一个函数 $V_p(a)$ 的常数倍. (提示: 首先确定某 p 适合 $V(p) > 0$.)
5. 应用习题 2 的公式证明: 对于任意正整数 a 和 $b, ab = (a,b)[a,b]$.(对于第二种证明, 参看 1.7 节的习题 14(c).)
6. 证明素数的个数是无限的 (欧几里得). (提示: 若 $p_1p_2, \cdots, p_n$ 为 n 个素数, 则这些素数没有一个能整除整数 $p_1p_2\cdots p_n + 1$.)
*7. 定义函数 $e(n)$(n 为任意正整数) 为 n 的素因子分解中出现的指数的最大公因子. 证明
   (a) 对于 $\mathbf{Z}$ 中已知的 r 和 n, 存在整数 x 适合 $x^r = n$ 当且仅当 $r|e(n)$;
   (b) $e(n^r) = r \cdot e(n)$; (c) 若 $e(m) = e(n) = d$, 则 $d|e(mn)$.
8. 如果正整数之积 mn 为二次幂, 并且 $(m,n) = 1$, 证明 m 和 n 都为二次幂.
*9. 假定整数 x, y 和 z 没有 ± 1 以外的公因子, 以 x, y 和 z 为边长的直角三角形可以按以下方式找到.
   (a) 如果 $x^2 + y^2 = z^2$, 证明 x 和 y 不能都是奇数.
   (b) 如果 y 是偶数, 应用习题 8 证明: $y = 2mn$, 其中 m 和 n 为整数, $x = m^2 - n^2, z = m^2 + n^2$. (提示: 分解因子 $z^2 - x^2$, 并证明 $(z+x, z-x) = 2$.)

1.9 同 余 式

在确定一天的时间时, 通常只计算到 12 小时, 超过 12 小时后重新开始计算. 这种抛弃固定数 12 的倍数的简单想法是"同余"这个算术概念的基础. 如果两个整数只差 12 的整数倍, 我们就称它们对模 12 同余, 例如, 7 和 19 是同余的, 我们把它记作

$$7 \equiv 19 (\text{mod } 12).$$

定义 $a \equiv b(\text{mod } m)$ 成立当且仅当 $m|(a-b)$.

我们也可以说: $a \equiv b(\text{mod } m)$ 的意思是差 $a - b$ 在 m 的所有倍数的集合中. 另外还可以根据下述事实来定义: 每个整数 a 除以 m 剩下唯一的余数 (1.7 节的推论 1). 我们把这种定义叙述如下.

定理 13 两个整数 a 和 b 对模 m 同余当且仅当它们除以 $|m|$ 时剩下相同的余数.

因为 $a \equiv b(\text{mod } m)$ 当且仅当 $a \equiv b(\text{mod } -m)$, 所以只须对于 $m > 0$ 的情形证明这个定理.

证明 按照我们的定义, 首先假定 $a \equiv b(\bmod m)$, 那么 $a-b=cm$ 是 m 的倍数. b 除以 m 剩下余数 $b-qm=r$, 其中 $0 \leqslant r < m$. 则

$$a=b+cm=(qm+r)+cm=(q+c)m+r.$$

这个方程表明, r 是 a 除以 m 的唯一的余数, 因此 a 和 b 具有相同的余数.

反过来, 假设 $a=qm+r, b=q'm+r$, 它们具有同一个余数 r. 那么 $a-b=(q-q')m$ 可被 m 整除, 所以 $a \equiv b(\bmod m)$ 证毕

固定模 m 的同余关系具有下列性质, 即相等关系的定律(1.2 节) 的再现, 对一切整数 a,b 和 c 有

$$\left.\begin{array}{ll}\text{自反律} & a \equiv a \\ \text{对称律} & a \equiv b \text{ 意味着 } b \equiv a \\ \text{传递律} & a \equiv b \text{ 和 } b \equiv c \text{ 意味着 } a \equiv c\end{array}\right\} (\bmod m).$$

这些定律中的每一个都可以用同余的定义来证明. 按同余的定义, 对称律要求由 $m|(a-b)$ 推出 $m|(b-a)$. 这里假设条件是 $a-b=dm$, 把它写成 $b-a=(-d)m$, 便得出结论 $m|(b-a)$.

固定模 m 的同余关系还具有“代换性质”, 这也是相等关系的性质之一, 即: 同余整数之和同余, 而且同余整数之积同余.

定理 14 *如果 $a \equiv b(\bmod m)$, 那么对一切整数 x, 有*

$$a+x \equiv b+x, \quad ax \equiv bx, \quad -a \equiv -b \quad (\bmod m).$$

这里还用定义证明. 于是假设条件变成 $a-b=km$(对某个整数 k), 故有

$$m|(a+x-b-x), \quad m|(ax-bx), \quad m|(-a+b).$$

由此我们便可以导出结论.

对于方程成立的消去律对于同余式不一定成立. 例如, 由 $2 \cdot 7 \equiv 2 \cdot 1(\bmod 12)$. 不能推出 $7 \equiv 1(\bmod 12)$. 之所以不能这样推断, 是因为被消去的 2 是模的一个因子. 对于同余, 最好也只能得到修改的消去律:

定理 15 *当 c 与 m 互素时,*

$$\text{由 } ca \equiv cb(\bmod m) \text{ 可推出 } a \equiv b(\bmod m).$$

证明 根据定义, 假设条件表明 $m|(ca-cd)$, 或 $m|c(a-b)$. 但是已假定 m 与这个乘积的第一个因子 c 互素, 因此由定理 10 得到 m 整除第二个因子 $a-b$. 这意味着 $a \equiv b(\bmod m)$, 如断言所述.

线性方程的讨论可以扩展到同余式上.

定理 16 *如果 c 与 m 互素, 那么同余式*

$$cx \equiv b \quad (\bmod m)$$

有整数解 x. 任意两个解 x_1 和 x_2 对模 m 同余.

证明　根据假设 $(c,m)=1$, 对适当的整数 s 和 t, 有 $1=sc+tm$. 两边乘以 $b, b=bsc+btm$. 这里最后一项是 m 的倍数, 因此 $b\equiv(bs)c(\mathrm{mod}\ m)$. 这就表明 $x=bs$ 是同余式 $b\equiv xc$ 所要求的解.

另一方面, 因为同余关系满足传递律和对称律, 所以这个同余式的两个解 x_1 和 x_2 必满足 $cx_1\equiv cx_2(\mathrm{mod}\ m)$. 因为已假设 c 与 m 互素, 所以我们可以像定理 15 那样消去这里的 c, 而得到所需要的结论 $x_1\equiv x_2(\mathrm{mod}\ m)$.　证毕

当模 m 为素数时, 出现重要的特殊情形. 在这种情形下, 不能被 m 整除的一切整数都与 m 互素. 由此得出

推论　*如果 p 为素数, 并且 $c\not\equiv 0(\mathrm{mod}\ p)$, 那么 $cx\equiv b(\mathrm{mod}\ p)$ 有模 p 的唯一解.*

也可以解联立同余式.

定理 17　*如果模 m_1 和 m_2 互素, 那么同余式*

$$x\equiv b_1 \qquad (\mathrm{mod}\ m_1), x\equiv b_2 \qquad (\mathrm{mod}\ m_2) \tag{17}$$

有公共解 x. 任意两个解对模 m_1m_2 同余.

证明　对任意整数 y, $x=b_1+ym_1$ 是第一个同余式的解. 这样的 x 满足第二个同余式当且仅当 $b_1+ym_1\equiv b_2(\mathrm{mod}\ m_2)$ 或 $ym_1\equiv b_2-b_1(\mathrm{mod}\ m_2)$. 因为 m_1 与 m_2 互素, 根据定理 16, 这个同余式对 y 可解,

另一方面, 假设 x 和 x' 是已知联立同余式 (17) 的两个解. 那么 $x-x'\equiv 0(\mathrm{mod}\ m_1)$ 和 $(\mathrm{mod}\ m_2)$. 因为 m_1 与 m_2 互素, 这意味着差 $x-x'$ 可被乘积模 m_1m_2 整除, 因此 $x\equiv x'(\mathrm{mod}\ m_1m_2)$.　证毕

上面同样的方法应用于形为

$$a_ix\equiv b_i \quad (\mathrm{mod}\ m_i)$$

的两个或多个同余式, 其中 $(a_i,m_i)=1$, 并且各个不同的模两两互素.

定理 18　(费马 (Fermat)) *如果 a 为整数, 并且 p 为素数, 那么*

$$a^p\equiv a \qquad (\mathrm{mod}\ p).$$

证明　对固定的 p, 设 $P(n)$ 为命题：$n^p\equiv n(\mathrm{mod}\ p)$. 那么 $P(0)$ 和 $P(1)$ 显然正确. 在 $(n+1)^p$ 的二项展开式 (9) 中, 除第一个和最后一个系数外, 每个系数都能被 p 整除, 因此 $(n+1)^p\equiv n^p+1(\mathrm{mod}\ p)$, 由 $P(n)$ 推出 $(n+1)^p\equiv n+1(\mathrm{mod}\ p)$, 这就是命题 $P(n+1)$.

习　题

1. 解下列同余式：

(a) $3x\equiv 2(\mathrm{mod}\ 5)$,　　(b) $7x\equiv 4(\mathrm{mod}\ 10)$,

(c) $243x+17\equiv 101(\mathrm{mod}\ 725)$,　　(d) $4x+3\equiv 4(\mathrm{mod}\ 5)$,

(e) $6x+3\equiv 4(\mathrm{mod}\ 10)$,　　(f) $6x+3\equiv 1(\mathrm{mod}\ 10)$.

2. 证明：关系 $a\equiv b(\mathrm{mod}\ m)$ 满足自反律和传递律.

3. 直接证明：由 $a\equiv b(\mathrm{mod}\ m)$ 和 $c\equiv d(\mathrm{mod}\ m)$ 可推出 $a+c\equiv b+d(\mathrm{mod}\ m)$ 和 $ac\equiv bd(\mathrm{mod}\ m)$.

*4. (a) 证明：同余式 $ax\equiv b(\mathrm{mod}\ m)$ 有解当且仅当 $(a,m)|b$.

(b) 证明：如果 $(a,m)|b$, 那么同余式恰有 (a,m) 个对模 m 不同余的解. (提示：用 (a,m) 除 a,b 和 m.)

5. 如果 m 为整数, 证明：$m^2\equiv 0,1$ 或 $4(\mathrm{mod}\ 8)$.

6. 证明：$x^2\equiv 35(\mathrm{mod}\ 100)$ 无解.

*7. 证明：如果 $x^2\equiv n(\mathrm{mod}\ 65)$ 有解, 那么 $x^2\equiv -n(\mathrm{mod}\ 65)$ 也有解.

8. 如果 x 为不能被 3 整除的奇数, 证明：$x^2\equiv 1(\mathrm{mod}\ 24)$.

*9. (a) 列表指出 $25\sim 40$ 中所有可表为四个或不多于四个平方和的整数 (这个结果实际上对于一切正整数都成立).

(b) 证明：满足 $m\equiv 7(\mathrm{mod}\ 8)$ 的任何整数不能表为三个平方之和. (提示：应用习题 5.)

10. 解联立同余式：

(a) $x\equiv 2(\mathrm{mod}\ 5)$, $2x\equiv 1(\mathrm{mod}\ 8)$;　　(b) $3x\equiv 2(\mathrm{mod}\ 5)$, $2x\equiv 1(\mathrm{mod}\ 3)$.

11. 在一个荒岛上, 五个人和一只猴子采了整整一天的椰子, 然后去睡觉. 第一个人醒来, 决定拿走他的那份椰子. 他把椰子平均分成五份, 正好多余一个, 分给猴子, 他藏起自己的那一份后, 就去睡觉了. 后来, 第二个人醒来也在剩下的一堆椰子中取出他那五分之一, 并把多余的一个分给猴子. 其余三个人也都照样做一遍. 求出原来摘下的一堆椰子的最小数目. (提示：列出同余式并用 -4 去试.)

*12. 用归纳法证明：定理 17 可以推广到 n 个同余式, 其模两两互素.

*13. 证明：如果 $(m_1,m_2)=(a_1,m_1)=(a_2,m_2)=1$, 那么联立同余式 $a_ix\equiv b_i(\mathrm{mod}\ m_i)(i=1,2)$ 有公共解, 并且任意两个解对模 m_1m_2 同余.

*14. 把习题 13 推广到 n 个联立同余式.

15. 对于什么样的正整数 m, 命题"如果 $x^2\equiv 0(\mathrm{mod}\ m)$, 那么也有 $x\equiv 0(\mathrm{mod}\ m)$"是正确的?

16. 如果 a 和 b 为整数, 并且 p 为素数, 证明：$(a+b)^p\equiv a^p+b^p(\mathrm{mod}\ p)$.

1.10 环 $\mathbf{Z}_n$

古代, 人们已区分了"偶"数 $2,4,6,\cdots$ 和"奇"数 $1,3,5\cdots$. 下面计算偶数和奇数的法则是大家熟悉的：

$$\text{偶数}+\text{偶数}=\text{奇数}+\text{奇数}=\text{偶数},$$
$$\text{偶数}+\text{奇数}=\text{奇数},$$

$$
\begin{aligned}
&\text{偶数}\cdot\text{偶数} = \text{偶数}\cdot\text{奇数} = \text{偶数},\\
&\text{奇数}\cdot\text{奇数} = \text{奇数}.
\end{aligned} \tag{18}
$$

这些恒等式定义一个新的整环 $\mathbf{Z}_2$, 它仅由两个元素 0("偶数") 和 1("奇数") 组成, 并且有加法表和乘法表:

$$
\begin{aligned}
&0+0=1+1=0, \qquad 0+1=1+0=1,\\
&0\times 0=0\times 1=1\times 0=0, \quad 1\times 1=1.
\end{aligned}
$$

我们现在要指出, 类似的构造可用于对任意模 n 的全体剩余 $0,1,2,\cdots,n-1$. 两个这样的剩余相加 (或相乘), 可以先简单进行普通意义下 (即在 $\mathbf{Z}$ 中) 的加法 (或乘法), 然后将所得结果取模 n 的剩余. 在 $n=5$ 的情形中, 其加法表和乘法表是:

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

×	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

对于每个 n 所得到的系统都具有 1.1 节的性质 (i)~(viii). 也就是说, 我们有

定理 19 *在加法和乘法之下, 对任意固定的模 $n\geqslant 2$, 整数 $0,1,\cdots,n-1$ 的集合组成一个交换环 $\mathbf{Z}_n$.*

证明 在上一节里我们看到, 关系 $x\equiv y(\bmod n)$ 同普通的相等关系一样, 满足自反律、对称律和传递律. 事实上, 根据定理 14, 由 $a\equiv b(\bmod n)$ 和 $c\equiv d(\bmod n)$ 推出

$$
a+c\equiv b+d(\bmod n), ac\equiv bd(\bmod n). \tag{19}
$$

也就是说, 倘若 $\mathbf{Z}$ 中的"相等"重新解释为"对模 n 同余", 则公设 (i) 和 (ii) 成立. 再有, $\mathbf{Z}$ 中的 0 和 1 在 $\mathbf{Z}_n$ 中分别起加法单位元素和乘法单位元素的作用, 而 $n-k$ 是 k 对模 n 的加法逆元素.

剩下来证明公设 (iii)~(v). 考虑分配律, 因为对任意整数, 有 $a(b+c)=ab+ac$, 所以当取模 n 的剩余时, 根据 (19), 我们必有 $a(b+c)\equiv ab+ac(\bmod n)$. 这就是 $\mathbf{Z}_n$ 中的分配律. 交换律和结合律的证明也完全类似. 证毕

与整环定义唯一不相一致的公设是乘法消去律. 根据定理 1, 这个定律等价于断语: $\mathbf{Z}_n$ 中无零因子, 即由 $ab=0$ 推出 $a=0$ 或 $b=0$. 这些方程在 $\mathbf{Z}_n$ 中表示普通整数的同余式, 所以定律表述为: 由 $ab\equiv 0(\bmod n)$ 推出 $a\equiv 0(\bmod n)$ 或 $b\equiv 0(\bmod n)$. 这等价于断语: 由 $n|ab$ 推出 $n|a$ 或 $n|b$. 如果 n 是素数, 这是正确的 (定理 9). 如果 n 不是素数, n 具有非平凡的因子分解 $n=ab$, 则 $n|ab$, 显然, $n|a$ 和 $n|b$ 都不成立, 因此 $\mathbf{Z}_n$ 有零因子. 这就证明了

定理 20 模 n 整数环 $\mathbf{Z}_n$ 是整环当且仅当 n 是素数.

还有其他更系统的方法构造模 n 整数的代数. 用等式代替同余式的方法本质上意味着: 把所有用 n 去除而剩下同样余数的整数归在一组, 产生一个新的"数". 每个这样的整数组称为"剩余类". 对于模 5, 有五个这样的类对应着可能的余数 $0, 1, 2, 3$ 和 4, 其中的一些是

$$1_5 = \{\cdots, -14, -9, -4, 1, 6, 11, 16, \cdots\},$$
$$2_5 = \{\cdots, -13, -8, -3, 2, 7, 12, 17, \cdots\},$$
$$3_5 = \{\cdots, -12, -7, -2, 3, 8, 13, 18, \cdots, \}.$$

对于任意模 n 由余数 $r(0 \leqslant r < n)$ 确定的剩余类 r_n, 是由所有用 n 去除而剩下余数 r 的整数 a 组成. 每个整数属于一个且仅属于一个剩余类, 而且两个整数属于同一个剩余类当且仅当它们同余 (定理 13). 模 n 有 n 个剩余类: $0_n, 1_n, \cdots, (n-1)_n$.

$\mathbf{Z}_n$ 的代数运算可以直接在这些类上进行. 假定两个剩余 r 和 s 在 $\mathbf{Z}_n$ 中给出剩余 t 作为它们的和, $r + s \equiv t (\text{mod } n)$. 如果我们用相应类中的任何其他元素来代替剩余 r 和 s, 便可得到上面的回答. 若 a 在 r_n 中, b 在 s_n 中, 则 $a + b$ 在属于和 t 的类 t_n 中, 这是因为, $a \equiv r$ 和 $b \equiv s$ 得出 $a + b \equiv r + s \equiv t (\text{mod } n)$. 一般地, 代数 $\mathbf{Z}_n$ 可以定义为这些剩余类的代数: 两个剩余类相加 (或相乘), 可在这两个类中任意选择代表元素 a 和 b, 并求出含有这两个代表元素的和 (或积) 的剩余类. 如果 a_n 表示包含 a 的剩余类, 这个法则可表述为

$$(a + b)_n = a_n + b_n, \quad (ab)_n = a_n b_n. \tag{20}$$

例如, 上面列出的剩余类中, 和 $1_5 + 2_5 = 3_5$ 可以这样求出: 在剩余类 1_5 和 2_5 中任意选出代表元素 6 和 (-13), 把它们相加得 -7, 而 -7 在和类 3_5 中. 其他选法 $(-9) + (-3) = -12$, $11 + 7 = 18$, $(-14) + 17 = 3$, 它们都给出同一个和 3_5.

我们利用剩余定义的剩余类也可以用 6.13 节中讨论的一般方法通过同余式直接定义.

习 题

1. 构造 $\mathbf{Z}_3$ 和 $\mathbf{Z}_4$ 的加法表和乘法表.
2. 在 $\mathbf{Z}_7$ 中计算: $(3 \cdot 4) \cdot 5$, $3 \cdot (4 \cdot 5)$, $3 \cdot (4 + 5)$, $3 \cdot 4 + 3 \cdot 5$.
3. 求出 $\mathbf{Z}_{26}$ 和 $\mathbf{Z}_{24}$ 的全部零因子.
4. 对于 4_8 中的 x 和 y, 确定所有和 $x + y$ 的确切集合及所有积 xy 的确切集合. 它们与集合 $4_8 + 4_8$ 及 $4_8 \cdot 4_8$ 有何关系?
5. 像证明定理 19 那样证明剩余类加法的结合律.

6. 对于实数 x 和 y 设 $x \equiv y(\text{mod } 2\pi)$ 表示 $x = y + 2n\pi$, 对某整数 n. 证明：剩余类的加法可以像 (20) 式那样定义, 而剩余类的乘法则不能这样定义.

*7. 证明：在 $\mathbf{Z}_n$ 中, 不是单位的任何元素 c 是零因子.

*8. (a) 列出 $\mathbf{Z}_{15}$ 的单位.

(b) 证明：若 $n = 2m + 1$ 是奇数, 则 $\mathbf{Z}_n$ 的单位的个数是偶数.

*9. 证明：k 是 $\mathbf{Z}_n$ 的单位当且仅当在 $\mathbf{Z}$ 中 $(k, n) = 1$.

1.11 集合 · 函数 · 关系

这一节我们暂时简短地讨论一下集合、函数、二元运算和关系等基本概念.

集合是一些数学对象完全任意的集体. 例如, 所有奇数的集合, 或平面上所有到两定点距离相等的点的集合. 如果 A 是集合, 则我们记 $x \in A$ 表示对象 x 是集合 A 的元素, 当 x 不是 A 的元素时, 记作 $x \notin A$. 有限集合可以通过列出它的所有元素来确定, 例如, $\{0, 2, 4\}$ 表示一个集合, 它仅有的元素是 0, 2 和 4, 更一般地, 任何集合由它的元素来确定. 在这种意义下. 两个集合 A 和 B 相等 (相同) 当且仅当它们具有相同的元素. 这个原则 (称为*外延性公理*) 也可用符号表达为 $A = B$, 其意思是, 对一切 $x, x \in A$ 当且仅当 $x \in B$. 这样得到的集合相等关系, 显然像 1.2 节中对任意相等关系要求的那样, 满足自反律、对称律和传递律.

集合 S 称为集合 A 的子集当且仅当 S 的每个元素 x 也在 A 中, 符号 $S \subset A$ 表示 S 是 A 的子集. 如果 $T \subset S$ 和 $S \subset A$ 两者都成立, 那么显然有 $T \subset A$, 因此关系 “$\subset$” 满足传递律. 集合相等的条件也可变成：$A = B$ 当且仅当 $A \subset B$ 和 $B \subset A$ 两者都成立. 此外, 空集$\varnothing$(没有元素的集合) 是每个集合的子集.

从任意集合出发, 比如全体整数的集合, 我们可以选出各种不同的子集：所有正整数的集合, 所有正奇数的集合, 所有大于 18 的整数的集合, 等等. 这些例子说明一个原则：任何性质都可确定一个子集；更确切地说, 已知任意集合 A 和性质 P, 我们可以构成子集

$$S = \{x | x \in A, \text{并且} x \text{具有性质} P\}, \tag{21}$$

它是由 A 中具有性质 P 的所有元素组成.

一般地, 如果 A 和 B 都是集合, 则关于 A 到 B 的*函数*$\phi : A \to B$ 是这样规定的, 它对 A 中每个元素 a 给定 B 中的一个元素 $a\phi$. 我们把它记作 $a \longmapsto a\phi$. 例如 $x \longmapsto x^2$ 是关于所有有理数的集合 $A = \mathbf{Q}$ 到所有非负有理数的集合 B 的函数ϕ(它也可看作函数 $\phi : \mathbf{Q} \to \mathbf{Q}$). 还有 “加一” 运算 $n \longmapsto n + 1$, 它把每个整数 n 传送到 $n + 1$, 因此它是一个函数 $\phi : \mathbf{Z} \to \mathbf{Z}$. 在任意有序整环 D 中, 取绝对值的运算是关于集合 D 到 D 中非负元素集合的一个函数. “取负” 运算 $a \longmapsto -a$ 是关于 D 到 D 的另一个函数.

关系 $a \longmapsto a\phi$ 有时写成 $a \longmapsto \phi a$ 或 $a \longmapsto \phi(a)$，这里函数的符号 ϕ 写在前面. 函数 $\phi: A \to B$ 也称为由 A 到 B 的映射、变换或对应. 集合 A 称为函数 ϕ 的定义域, 而 B 是函数 ϕ 的变换的取值域. 例如, 通常的电话拨号盘

ABC	DEF	GHI	JKL	MNO	PRS	TUV	WXY	Z
2	3	4	5	6	7	8	9	0

定义了关于 25 个字母 (字母表略去 Q) 的集合 A 到 10 个数字的集合 $\{0,1,2,\cdots,9\}$ 的函数.

函数 $\phi: A \to B$ 的像(或"值域") 是所有函数"值"的集合, 即所有 $a\phi$(a 在 A 中) 的集合, 像是取值域 B 的子集, 而不一定是整个 B. 例如电话拨号盘函数的像是略去了 1 的一个子集 $\{0, 2, \cdots, 9\}$.

函数 $\phi: A \to B$, 当 B 的每个元素 b 是函数的像时, 也就是说, 当像是整个取值域时, 称 ϕ 是满射(映上). 例如, 整数取绝对值 $a \longmapsto |a|$ 是关于 $\mathbf{Z} \to \mathbf{Z}$ 的函数, 但它不是满射, 因为像是所有非负整数 $\mathbf{N} \subset \mathbf{Z}$, 它是 $\mathbf{Z}$ 的真子集. 但是法则 $a \longmapsto |a|$ 也定义了 $\mathbf{Z} \to \mathbf{N}$ 的函数, 而它是满射. 为决定函数是否是映上的, 我们必须知道预定的取值域.

函数 $\phi: A \to B$, 当 A 的不同元素总有不同的像时, 换句话说, 当由 $a\phi = a'\phi$ 总可推出 $a = a'$ 时, 则称 ϕ 是单射 (一一映入). 例如, $x \longmapsto 2x$ 是 $\mathbf{Z} \to \mathbf{Z}$ 的一个单射(但不是满射).

函数 $\phi: A \to B$, 当它既是单射又是满射, 即当对每个元素 $b \in B$, 有一个且仅有一个 $a \in A$ 具有像 b, 使 $a\phi = b$, 则 ϕ 是双射 (一一映上). 例如, $n \longmapsto n+1$ 是 $\mathbf{Z} \to \mathbf{Z}$ 的双射. 还有, 对任意整环 $D, a \longmapsto a$ 是 $D \to D$ 的双射. 双射 $\phi: A \to B$ 也称为 (A 到 B 上的)一一对应, 而不是单射的对应称为多一对应.

二元运算 数对的运算出现在很多方面 —— 两个整数的加法, $\mathbf{Z}_n$ 中两个剩余类的加法, 两个实数的乘法, 一个整数减去另一个整数的减法, 等等. 在这种情况下, 我们称这些运算为二元运算. 一般地, 元素 $a, b, c\cdots$ 的集合 S 上的二元运算"$\circ$"是这样规定: 它对 S 中每个有序元素对 a 和 b 给出同一集合 S 中唯一确定的第三个元素 $c = a \circ b$. 这里我们用"唯一"表示代换性质

$$\text{由 } a = a' \text{ 和 } b = b' \text{ 推出 } a \circ b = a' \circ b', \tag{22}$$

同交换环的唯一性公设中所说的一样.

为方便起见, 把所有有序元素对 (a, b)(其中 $a \in S, b \in T$) 的集合记作 $S \times T$, 这称为 S 和 T 的笛卡儿(Cartersian) 积(或简称"积"). 我们又把集合同自身的积 $S \times S$ 记作 S^2, 那么二元运算同函数 $\circ: S^2 \to S$ 一样.

两个已知整数之间可以有多种关系, 例如 “$a=b$”, “$a<b$”, “$a\equiv b(\mathrm{mod}\ 7)$”, 或 “$a|b$”. 上述每个语句都表示 a 和 b 之间的某个 “二元关系”. 我们可以容易地叙述其他类型的数学对象之间的许多别的关系, 也有像人与人之间的 “是 $\cdots$ 的兄弟” 这样一类的非数学的关系. 为一般地讨论关系, 我们引进符号 R 来表示任何关系 (“R” 代表 “$<$”, “$\equiv$” 或 “$|$”, 等等). 形式上, 如果已知集合 S 中的两个元素 a 和 b, 不是 a 与 b 有关系 R(记号为 aRb), 就是 a 与 b 没有关系 R(记号为 $aR'b$), 那么 “R” 就表示集合 S 上的二元关系.

数学中特别重要的是考虑像同余和相等那样在集合 S 上满足下列定律的关系 R:

自反律 aRa, 对 S 中一切 a;

对称律 由 aRb 可推出 bRa, 对 S 中一切 a,b;

传递律 由 aRb 和 bRc 可推出 aRc, 对 S 中一切 a,b,c.

满足自反律、对称律和传递律的关系称为*等价关系*. 例如, 平面上三角形之间的全等关系就是这样的等价关系.

习 题

1. 下列整数 a 和 b 的二元运算 $a\circ b$ 中, 哪些满足结合律? 哪些满足交换律?

$$a-b,\quad a^2+b^2,\quad 2(a+b),\quad -a-b.$$

2. “自反律”、“对称律” 和 “传递律” 三种性质中, 哪一种适用于下列整数 a 和 b 之间的所有关系?

$$a\leqslant b,\quad a<b,\quad a|b,\quad a^2+a=b^2+b,\quad a<|b|.$$

3. 上面三种性质对下列人的分类关系是否适合: “是 $\cdots\cdots$ 的父亲”, “是 $\cdots\cdots$ 的兄弟”, “是 $\cdots\cdots$ 的朋友”, “是 $\cdots\cdots$ 的叔叔”, “是 $\cdots\cdots$ 的子孙”. 如果这些关系被限定只用于一切男人的分类, 那么你的回答中哪些会有变化?

*4. 关系 “是 $\cdots\cdots$ 的叔叔” 与关系 “是 $\cdots\cdots$ 的兄弟” 和 “是 $\cdots\cdots$ 的父亲” 是怎样联系的? 你能叙述一下由两个已知关系做出新关系的类似的一般法则吗?

5. 如果关系 R 由 aRb 和 bRc 推出 cRa, 则称 R 为循环的. 证明: 关系 R 是自反的和循环的当且仅当它满足自反律、对称律和传递律.

*6. 下面由关系 R 的对称律和传递律推出自反律的 “证明” 其错误是什么?
“根据对称律, aRb 推出 bRa; 根据传递律, aRb 和 bRa 推出 aRa.”

7. 下列法则中, 每一个都定义一个函数 $f:\mathbf{Z}\to\mathbf{Z}$. 对每种情况详细说明其像, 以及函数是否是单射.

(a) $a\longmapsto |a|+1$, (b) $a\longmapsto a^2$,

(c) $a\longmapsto 2a+5$, (d) $a\longmapsto \mathrm{g.c.d.}(a,\ 6)$.

8. 用正整数的集合 $\mathbf{Z}^+$ 代替 $\mathbf{Z}$, 做习题 7.

9. 对什么样的整数 n, 函数 $x \longmapsto 6x+7$ 在 $\mathbf{Z}_n$ 上是双射? 对什么样的 n, 函数 $x \longmapsto 6x+7$ 在 $\mathbf{Z}_n$ 上是满射?

10. 证明: 集合 S 上任何关系 R 都可以看作函数 $f: S^2 \to \{0,1\}$.

1.12 同构与自同构

近世代数最重要的概念之一是同构的概念. 我们现在对交换环如下定义这个概念:

定义 *两个交换环 R 和 R' 之间的同构是 R 的元素 a 与 R' 的元素 a' 的一一对应 $a \leftrightarrow a'$, 并对所有元素 a 和 b 满足条件*

$$(a+b)' = a' + b', \qquad (ab)' = a'b'. \tag{23}$$

如果两个环 R 和 R' 之间存在这样的对应, 则称它们是同构的.

基于规律 (23), 我们可以说, 同构 $a \leftrightarrow a'$ "保持和与积". 粗略地说, 两个交换环当它们的元素仅仅区别于记号时, 它们是同构的. 一个恰当的例子是"偶数"和"奇数"的代数 (像 1.10 节所讨论的那样) 同整环 $\mathbf{Z}_2$ 比较. 一一对应

$$\text{偶数} \leftrightarrow 0 \qquad \text{奇数} \leftrightarrow 1$$

是这两个整环之间的同构, 这是因为它们的对应元素是按照相同的法则 (参看 (18) 式) 相加和相乘的.

许多整环具有同它们自身的同构, 这样的同构是很重要的, 它称为*自同构*. 类似于几何图形中的对称性 (参看 6.1 节). 例如, 考虑整环 $\mathbf{Z}[\sqrt{2}]$, 在 1.1 节中它表示所有数 $m+n\sqrt{2}$ 的集合, 其中 m 和 n 在整数环 $\mathbf{Z}$ 中, 在非平凡的对应 $m+n\sqrt{2} \leftrightarrow m-n\sqrt{2}$ 之下, $\mathbf{Z}[\sqrt{2}]$ 与它自身同构. 这种对应是同构, 因为对任意 $a = m+n\sqrt{2}$ 和 $b = m_1+n_1\sqrt{2}$, 有

$$\begin{aligned}
(ab)' &= [(m+n\sqrt{2})(m_1+n_1\sqrt{2})]' \\
&= [(mm_1+2nn_1)+(mn_1+m_1n)\sqrt{2}]' \\
&= (mm_1+2nn_1)-(mn_1+m_1n)\sqrt{2}, \\
a'b' &= (m-n\sqrt{2})(m_1-n_1\sqrt{2}) \\
&= (mm_1+2nn_1)-(mn_1+m_1n)\sqrt{2}.
\end{aligned}$$

类似地有

$$(a+b)' = a' + b'.$$

任何同构 $a \leftrightarrow a'$ 不仅保持和与积, 而且保持差. 根据定义, $a-b$ 是方程 $b+x=a$ 的解, 所以 $b+(a-b)=a$. 因为对应保持和, 所以 $b'+(a-b)'=a'$, 这就是说, $(a-b)'$

是方程 $b'+x=a'$ 的 (唯一解), 或者说

$$(a-b)'=a'-b'.$$

另一个法则是

$$0'=0, \qquad 1'=1, \qquad (-a)'=-(a'). \tag{24}$$

总之, R 的零 (单位元素) 对应于 R' 的零 (单位元素).

后面我们将看到, 同构的概念普遍应用于代数系统. 我们甚至可以说, 抽象代数是研究代数系统那些在同构之下仍保持不变的性质.

在把整数系描述为有序整环 (其中每个正整数集合具有最小元素) 时我们曾要求: 对于所有的数学意义, 这些公设完整地描述了全体整数. 现在我们可以把它叙述得更确切 (将在 2.6 节中证明). 任意有序整环当它所包含的全体正元素集合是良序的, 它就同构于整数环 $\mathbf{Z}$. $\mathbf{Z}$ 的"精确到同构"的这个特征是最完全的了, 它可用我们已用过的任何形式的公设系得到, 因为一般地, 显然, 如果系统 S 满足这样的公设系, 而且 S' 是另一个同构于 S 的系统, 那么 S' 也必满足这些公设. 因此, 如果 S 满足加法交换律, 则对 S 中一切 a 和 $b, a+b=b+a$. 由于在已知同构之下, 它们的对应元素必相等, 所以 $(a+b)'=(b+a)'$. 因为同构保持和, 所以 $a'+b'=b'+a'$. 这就断言: 交换律在 R' 中也成立. 这种论证具有一般性, 可应用于我们的一切公设.

习 题

1. 证明: 性质 (24) 对任意同构都成立.
2. 设 $\mathbf{Z}[\sqrt{3}]$ 是所有数 $m+n\sqrt{3}(m,n\in\mathbf{Z})$ 的整环. 列出 $\mathbf{Z}[\sqrt{3}]$ 的一个非平凡自同构.
3. 证明: 对应 $m+n\sqrt{2}\leftrightarrow m+n\sqrt{3}$ 不是整环 $\mathbf{Z}[\sqrt{2}]$ 和 $\mathbf{Z}[\sqrt{3}]$ 之间的同构.
4. (a) 证明: 在任意同构之下, 满足方程 $x^2=1+1$ 的元素 x 必对应于满足方程 $y^2=1'+1'$ 的元素 y.

   (b) 利用 (a) 证明: $\mathbf{Z}[\sqrt{2}]$ 和 $\mathbf{Z}[\sqrt{3}]$ 之间不可能有同构.
5. 证明: 整数环 $\mathbf{Z}$ 没有非平凡的自同构.

*6. 证明: 只含有三个元素的整环必同构于 $\mathbf{Z}_3$.

7. 证明: 同构是等价关系 (即满足自反律、对称律和传递律).

第 2 章　有理数和域

2.1　域的定义

全体有理数组成的整环 **Q** 和全体实数组成的整环 **R**, 具有整数环 **Z** 所不具备的极重要的代数特征：在它们之中, 任何方程 $ax = b(a \neq 0)$ 是可解的. 具有这个性质的交换环称为域. 我们现在证明, 在任何交换环中, 如果所有非零元素有乘法逆, 那么除法是可能的, 并具有一些熟知的性质.

定义　*如果 F 是一个交换环, 并且对每个元素 $a \neq 0$, 它都包含一个"逆"元素 a^{-1}, 满足方程 $a^{-1}a = 1$, 那么 F 是域.*

容易证明, 在任何域中, 1.1 节的消去律 (ix) 是成立的. 这是因为如果 $c \neq 0$, 且 $ca = cb$, 那么

$$a = 1 \cdot a = (c^{-1}c)a = c^{-1}(ca) = c^{-1}(cb) = (c^{-1}c)b = 1b = b.$$

换句话说, 每个域是一个整环. 更一般地, 是域的子整环 (根据相同的理由). 相反地, 我们将在本节和下一节中指出, 任何整环能够按照唯一的最小路径被扩张成域. 我们通过把分数表为整数之商的标准表示法来说明扩张的方法.

定理 1　*在任何域中, 除法 (零除外) 是可能的而且是唯一确定的.*

证明　我们来证明, 在域 F 中, 对给定的 $a \neq 0$ 和 b, 方程 $ax = b$ 在 F 中有唯一解 x. 如果 $a \neq 0$, 可以用逆 a^{-1} 来构造一个元素 $x = a^{-1}b$, 代入方程可以验证它就是 $ax = b$ 的解. 这个解是唯一的, 因为根据前面证过的消去律, 当 $a \neq 0$ 时, 由 $ax = b$ 和 $ay = b$ 可以推出 $x = y$.　　证毕

我们用 $\dfrac{b}{a}$(a 除 b 所得的商) 表示 $ax = b$ 的解, 特别地, $\dfrac{1}{a} = a^{-1}$.

在被看作整环的域中, 1.2 节中列举的所有代数运算法则都成立. 通常的商的运算法则也可以由域的公设来证明.

定理 2　*在任何域中, 商遵循下列法则 (这里 $b \neq 0, d \neq 0$)：*

(i) $\dfrac{a}{b} = \dfrac{c}{d}$ 当且仅当 $ad = bc$,　　(ii) $\dfrac{a}{b} \pm \dfrac{c}{d} = \dfrac{ad \pm bc}{bd}$,

(iii) $\dfrac{a}{b} \cdot \dfrac{c}{d} = \dfrac{ac}{bd}$,　　(iv) $\dfrac{a}{b} + (-\dfrac{a}{b}) = 0$,

(v) $\dfrac{a}{b} \cdot \dfrac{b}{a} = 1$, 当 $\dfrac{a}{b} \neq 0$.

证明　(i) 假设条件 $\frac{a}{b}=\frac{c}{d}$ 意味着 $ab^{-1}=cd^{-1}$, 由此得

$$ad=a(b^{-1}b)d=cd^{-1}(bd)=cd^{-1}ab=bc.$$

反过来, 如果 $ad=bc$, 那么

$$\frac{a}{b}=b^{-1}a=b^{-1}add^{-1}=b^{-1}bcd^{-1}=cd^{-1}=\frac{c}{d},$$

即所要证的.

(ii) 注意到 $x=\frac{a}{b}$ 和 $y=\frac{c}{d}$, 分别表示 $bx=a$ 和 $dy=c$ 的解, 这些方程还可以组合成

$$dbx=da,\qquad bdy=bc,\qquad bd(x\pm y)=ad\pm bc.$$

于是 $x\pm y$ 是方程 $bdz=ad\pm bc$ 的唯一解 $z=\frac{ad\pm bc}{bd}$.

(iii) 如上所述, 方程 $bx=a$ 和 $dy=c$ 可组合成

$$(bd)(xy)=(bx)(dy)=ac,$$

因此

$$xy=\frac{ac}{bd}.$$

(iv) 在 (ii) 中用 $-\frac{a}{b}$ 代替 $\frac{c}{d}$, 我们有

$$\frac{a}{b}+\left(-\frac{a}{b}\right)=\frac{ab-ba}{b^2}=\frac{0}{b^2}=0(b^2)^{-1}=0.$$

(v) 在 (iii) 中用 $\frac{b}{a}$ 代替 $\frac{c}{d}$, 我们有 $\frac{a}{b}\cdot\frac{b}{a}=\frac{ab}{ba}$, 而 $\frac{ab}{ba}$ 是方程 $bax=ab$ 的唯一解. 显然, $x=1$ 满足这个方程, 因此 $\frac{ab}{ba}=1$.　　证毕

用类似于刚用过的那些论证, 可以证明下列其他熟悉的定律:

$$(bd)^{-1}=d^{-1}b^{-1},\qquad (-b)^{-1}=-(b^{-1}),\qquad \text{当}b,d\neq 0.\tag{1}$$

$$a\pm\frac{b}{c}=\frac{ac\pm b}{c},\qquad a\frac{b}{c}=\frac{ab}{c},\qquad \text{当}c\neq 0.\tag{2}$$

$$\frac{a}{b}\Big/\frac{c}{d}=\frac{ad}{bc},\qquad \frac{a}{b}\Big/c=\frac{a}{bc},\qquad \frac{a}{1}=a,\qquad \text{当 } b,c,d\neq 0.\tag{3}$$

$$-\frac{a}{b}=\frac{-a}{b}=\frac{a}{-b},\quad \frac{-a}{-b}=\frac{a}{b},\quad \text{当}b\neq 0.\tag{4}$$

其证明将留给读者作习题.

存在各种各样的域. 例如, 对任意素数 p, 1.10 节中所构造的整环 $\mathbf{Z}_p$ 是一个域. 这由 1.9 节定理 16 的推论可以得到. 再有, 如果我们假定全体实数构成一个域, 那么我们利用子域的概念可以容易地构造出其他域的例子.

定义 *如果一个给定的域 F 的子集在 F 中的加法和乘法运算之下构成一个域, 那么称这个子集为域 F 的子域.*

只要问题中的运算能够进行, 那么所有在 F 中成立的恒等式 (即交换律、结合律和分配律) 在 F 的任何子集中自然成立. 因此在检验 F 的子集 S 是否是子域时, 我们可以不管那些恒等式的公设, 而只须检验那些包含某个 "存在性" 的公设, 比如, 逆元素的存在性. 这就给出下面的结果:

定理 3 *如果域 F 的子集 S 包含着 F 中的零元素和单位元素, S 在加法和乘法之下是封闭的, S 中每个 a 在 S 中有它的负元素和它的逆元素 a^{-1}(假定 $a \neq 0$), 那么 S 是子域.*

现在用定理 3 可以证明, 所有形如 $a+b\sqrt{2}$ 的实数的集合是实数域的一个子域, 其中系数 a 和 b 是有理数. 这个子域通常记作 $\mathbf{Q}(\sqrt{2})$, 这里 $\mathbf{Q}$ 表示有理数域. 可以应用定理 3 是因为, $\mathbf{Q}(\sqrt{2})$ 中任意两个数的和是另一个同样形式的数, 类似地, 两个数的积是

$$(a+b\sqrt{2})(c+d\sqrt{2}) = (ac+2bd)+(bc+ad)\sqrt{2}.$$

再有, $\mathbf{Q}(\sqrt{2})$ 包含 $0=0+0\sqrt{2}$, $1=1+0\sqrt{2}$, 并且如果它包含 $a+b\sqrt{2}$, 则也包含

$$-(a+b\sqrt{2}) = -a-b\sqrt{2}.$$

最后, 任何非零元素的逆元素 $(a+b\sqrt{2})^{-1}$ 可以通过 "分母有理化" 来求出,

$$\frac{1}{a+b\sqrt{2}} = \frac{1}{a+b\sqrt{2}}\left(\frac{a-b\sqrt{2}}{a-b\sqrt{2}}\right) = \frac{a}{a^2-2b^2} - \frac{b}{a^2-2b^2}\sqrt{2}.$$

新的分母 a^2-2b^2 不会是零 (在 3.6 节中将给出证明), 求得的逆元素具有所要求的形式 $a'+b'\sqrt{2}$, 其中系数

$$a' = \frac{a}{a^2-2b^2}, \quad b' = -\frac{b}{a^2-2b^2}$$

是有理数. 我们容易验证, 这个逆元素确实满足方程

$$(a'+b'\sqrt{2})(a+b\sqrt{2}) = 1.$$

类似地, 所有实数 $a+b\sqrt[3]{5}+c\sqrt[3]{25}$ 的集合 $\mathbf{Q}(\sqrt[3]{5})$ 是一个域, 其中 a,b,c 是有理数. 几乎同 $\mathbf{Q}(\sqrt{2})$ 一样, 在这个集合中, 加法、减法和乘法可以进行, 这里用到这样

一个事实：$(\sqrt[3]{5})^3 = 5$ 是有理数. 最后, 由于方程式

$$(a + b\sqrt[3]{5} + c\sqrt[3]{25})(x + y\sqrt[3]{5} + z\sqrt[3]{25}) = 1 + 0\sqrt[3]{5} + 0\sqrt[3]{25}$$

等价于一个联立线性方程组, 而方程组总是能够解出 x, y和z, 除非 $a = b = c = 0$, 于是 $(a + b\sqrt[3]{5} + c\sqrt[3]{25})^{-1}$ 可以计算出来.

如果我们假定存在一个由所有复数 $a + b\mathrm{i}$ 构成的域, 其中 $\mathrm{i} = \sqrt{-1}$, a与b 是实数, 那么我们还可以构造其他子域. 二次方程

$$\omega^2 + \omega + 1 = 0$$

在复数中有根 $\omega = \dfrac{-1 + \sqrt{-3}}{2} = -\dfrac{1}{2} + \dfrac{\sqrt{3}}{2}\mathrm{i}$. (注意, 因为

$$\omega^3 - 1 = (\omega - 1)(\omega^2 + \omega + 1) = 0,$$

所以 ω 是一个“虚”的单位立方根!) 所有数 $a + b\omega$(a, b 为有理数) 构成复数域的一个子域 $\mathbf{Q}(\omega)$, 这是因为

$$(a + b\omega) + (c + d\omega) = (a + c) + (b + d)\omega,$$

$$\begin{aligned}(a + b\omega)(c + d\omega) &= ac + (bc + ad)\omega + bd\omega^2 \\ &= (ac - bd) + (bc + ad - bd)\omega,\end{aligned}$$

这里用了等式 $\omega^2 = -\omega - 1$ 来去掉 ω^2 项. 进一步, 对任意 $a + b\omega \neq 0$, 它在这个集合中有一个逆元素, 这是因为

$$(a + b\omega)\left[\frac{-(b - a + b\omega)}{a^2 - ab + b^2}\right] = \frac{a^2 - ab + b^2}{a^2 - ab + b^2} = 1.$$

这个逆元素中的分母 $a^2 - ab + b^2$ 决不会是零, 因为

$$a^2 - ab + b^2 = \frac{a^2 + b^2}{2} + \frac{(a - b)^2}{2}$$

一定是正的, 除非 $a = b = 0$.

习　题

1. 从域的公设出发证明公式 (1) 至公式 (4).
2. 在 $\mathbf{Z}_{11}$ 中, 对每个 $c \neq 0$, 列出 c^{-1} 的表.

3. 如果假定实数集合是一个域, 下列实数子集中哪一个是域?
   (a) 全体正整数.
   (b) 全体形如 $a+b\sqrt{3}$ 的数, 此处 a, b 是有理数.
   (c) 全体形如 $a+b\sqrt[3]{5}$ 的数, 此处 a, b 是有理数.
   (d) 全体不是整数的有理数.
   (e) 全体数 $a+b\sqrt{5}$, 此处 a, b 是有理数.
4. 证明: 在定理 3 中, 条件 "$0 \in S$ 和 $1 \in S$" 可用条件 "S至少包含两个元素" 代替. (提示: 考虑 $ax=a$.)
*5. 证明: 由 1.1 节中的公设 (i), (ii) 和 (iv)~ (vii) 以及下面的 (viii′), 可以推出定律 $a+b=b+a$.
   (viii′) 对 R 中每个 a, 方程 $a+x=0$ 和 $y+a=0$ 在 R 中有解 x 和 y.
6. 每个与域同构的整环本身是域吗? 为什么?
7. 证明: 有理数域 $\mathbf{Q}$ 的唯一的子域是 $\mathbf{Q}$ 本身.
8. 对于子整环, 叙述并证明类似于定理 3 的定理.
9. 证明: $\mathbf{Q}(\sqrt{2})$ 的子域或者是 $\mathbf{Q}$ 本身, 或者是整个域 $\mathbf{Q}(\sqrt{2})$.
10. 如果 S 和 S' 是给定的域 F 的两个子域, 证明 S 和 S' 公共元素的集合也是一个子域.
11. 你能叙述关于 $\mathbf{Z}$(以及 $\mathbf{Z}_n$) 的可能子整环的一般性定理吗?
*12. 构造四元素域的加法表和乘法表, 假定这四元素域满足 1+1=0(加法是模 2 的), 并且存在元素 x, 使得 $x^2=x+1$.
*13. 找出习题 12 的域的所有子域.

2.2 有理数域的构造

在第 1 章中假定了全体整数的良序整环 $\mathbf{Z}$ 的存在, 现在我们将严格地证明, 有理数域 $\mathbf{Q}$(有序的) 能够由 $\mathbf{Z}$ 构造出. 实际上, 更一般地, 我们将证明, 类似的构造可以应用到任何整环上去.

仅仅由全体整数不能构成域, 由整数构造有理数在本质上恰是构造了包含全体整数在内的域. 显然, 这个域还必须包含所有方程 $bx=a$ 的解, 其中系数 $a, b \neq 0$ 都是整数. 为了从这些方程的解抽象地构造 "有理数", 我们简单地引入某些新记号 (或称数偶)$r=(a, b)$, 每个记号代表一个方程 $bx=a$ 的解. 为此我们必须说明, 这些新记号完全像域中的商 $\dfrac{a}{b}$ 那样可以相加、相乘和相等 (定理 2 的 (i)~(iii)).

不管我们从整数环 $\mathbf{Z}$, 还是从其他一些整环 D 出发, 上述的说明是很有意义的. 这可以确切地描述如下:

定义 设 D 为任意整环. D 的商域 $Q(D)$ 是由所有数偶 (a, b) 组成, 其中 $a, b \in D$ 并且 $b \neq 0$. 这种数偶的 "相等" 由下面约定来确定:

$$(a, b)=(a', b') \text{ 当且仅当 } ab'=a'b, \tag{5}$$

而数偶的和与积分别由下列约定来确定：

$$(a,b)+(a',b')=(ab'+a'b,bb'), \tag{6}$$

$$(a,b)\cdot(a',b')=(aa',bb'). \tag{7}$$

注意, 因为 D 不包含"零因子"(1.2 节定理 1), 在 (6) 和 (7) 中的乘积 $bb'\neq 0$, 所以 $Q(D)$ 在加法和乘法之下是封闭的.

我们希望数偶之间的"同余"关系"≡"与相等关系一致. 由于这个关系不是正式的恒等关系 ((a,b)与(a',b') 恒等的意思是 $a=a', b=b'$), 所以我们必须证明, 这个同余具有 1.2 节中列举的相等的性质 (对于正式的恒等, 这些性质将是显然的). 首先我们通过直接的论证可以证明"≡"满足自反律、对称律和传递律. 其次, 和与积在同余意义下是唯一确定的. 例如, 由 $(a,b)\equiv(a',b')$ 可推出 $(a,b)+(a'',b'')\equiv(a',b')+(a'',b'')$. 上面结论中的每个和用 (6) 式那样的公式给出, 而且所得的这两个数偶在 (5) 式的意义下是同余的当且仅当

$$(ab''+a''b)b'b''=(a'b''+a''b')bb''.$$

而这个等式是由假设条件 $(a,b)\equiv(a',b')$(即$ab'=a'b$) 得出. 类似地, 对于乘积的唯一性断言也是成立的. 我们得出结论, 由 (5) 式定义的相等具有所要求的性质.

现在可以验证 $Q(D)$ 中的各种代数定律. 例如分配律, 根据定义 (6) 和定义 (7), 按照下面的方法我们可以一步一步地化简定律的每一边. 设 r, r'和r'' 是任意三个数偶,

$$\begin{array}{ll}
r(r'+r'') & rr'+rr'' \\
(a,b)[(a',b')+(a'',b'')] & (a,b)(a',b')+(a,b)(a'',b'') \\
(a,b)(a'b''+a''b',b'b'') & (aa',bb')+(aa'',bb'') \\
(aa'b''+aa''b',bb'b'') & (aa'bb''+aa''bb',bb'bb'').
\end{array}$$

最后一行的两边给出了在 (5) 的意义下相等的数偶, 这是因为右边与左边的差别只是在右边所有项中多出现一个非零因子 b, 在数偶中这样一个额外因子使数偶总保持相等, 即 $(bx,by)\equiv(x,y)$, 因为根据 (5) 式这个等式相当于恒等式 $bxy=byx$.

$Q(D)$ 中这个分配律的清楚的证明只作为例证. 同样, 我们可以直接运用 D 中的定义和定律证明结合律和交换律. 加法单位元素 (零) 是数偶 (0, 1), 因为

$$(0,1)+(a,b)=(0\cdot b+1\cdot a,1\cdot b)=(a,b).$$

同样, 消去律也成立, 并且数偶 (1, 1) 是乘法单位元素. (a,b) 的负元素是 $-(a,b)=(-a,b)$. 这就验证了 1.1 节中所列的关于整环的一切公设.

定理 4 对任意整环 D, 商域 $Q(D)$ 是一个域.

证明 剩下只须证明每个方程 $rx=1$(其中 $r\neq 0$) 在 $\boldsymbol{Q}(D)$ 中有一个解. 也就是说, 对每个 $r\neq 0$, 在 $Q(D)$ 中存在 r 的逆元素. 这是容易证明的. 更一般地, 任意方程

$$(a,b)(x,y)\equiv(c,d),\quad \text{其中}\quad (a,b)\not\equiv(0,1), \tag{8}$$

由 (3) 式给出一个解, 即

$$(x,y)=(bc,ad). \tag{8'}$$

这是因为, 把 x,y 的值直接代入方程后有

$$(a,b)(bc,ad)=(abc,bad).$$

又因为 $abcd=badc$, 所以 $(abc,bad)=(c,d)$. 假设条件 $(a,b)\not\equiv(0,1)$ 保证了 $a\neq 0$, 因此 (x,y) 的第二项 ad 不为零, 正如有理数定义所要求的那样. 证毕

我们现在希望证明, $Q(D)$ 实际上包含着原来的整环 D 作为它的子整环, 换句话说, $Q(D)$ 实际上是 D 的扩张. 严格说来, 这是不可能的, 因为数偶 (a,b) 不像 D 中那样的元素. 不过我们可以把每个 $a\in D$ 与 $(a,1)$ 联系起来, 在相等、加法和乘法之下, $(a,1)$ 具有的性质完全像 a 一样, 如下所示:

$$\begin{gathered}(a,1)+(b,1)=(a\cdot 1+b\cdot 1,1\cdot 1)=(a+b,1),\\ (a,1)\cdot(b,1)=(ab,1\cdot 1)=(ab,1),\\ (a,1)\equiv(b,1)\quad \text{当且仅当}\quad a=b.\end{gathered}$$

我们可以断定, 一一对应 $a\leftrightarrow(a,1)$ 是给定的整环 D 到域 $Q(D)=F$ 的子整环上的一个同构. 此外, 方程 (8) 和 (8′) 表明, 任何数偶 $r=(a,b)\in Q(D)$ 是方程 $(b,1)r=(a,1)$ 或者 $br=a$ 的解, 因此 $r=(a,b)$ 是商 $\frac{a}{b}$, 这就证明了

定理 5 任何整环 D 能够同构地嵌入域 $Q(D)$ 中, $Q(D)$ 的每个元素是 D 中两个元素的商.

特别地, 把定理 5 用到整数环 $\mathbf{Z}$ 上. 事实上在上述论证中始终想到 $D=\mathbf{Z}$ 这一特殊情形, 因此 $Q(D)=Q(\mathbf{Z})$ 是全体普通分数的集合. 所以我们有

推论 整数环 $\mathbf{Z}$ 可以作为子整环嵌入域 $\mathbf{Q}=Q(\mathbf{Z})$ 中, 域 $\mathbf{Q}$ 的每个元素是整数的商 a/b, 其中 $b\neq 0$.

我们现在指出, 有理数域 $\mathbf{Q}=Q(\mathbf{Z})$ 实际上已通过前面的论述被精确地表征出来 (精确到同构). 因为 $\mathbf{Z}$ 是由它的公设所定义 (仅精确到同构), 所以这像我们所希望的那样是完备的表征. 事实上, 我们将证明, 任何整环 D 都有类似的结果.

定理 6 设整环 D 作为子整环包含在任意一个域 F 中, 那么 F 中所有形为 $\frac{a}{b}$(其中 $a,b\in D, b\neq 0$) 的元素组成的集合是 F 的一个子域 S, 并且在对应 $\frac{a}{b}\leftrightarrow(a,b)$ 之下这个子域 S 与 $Q(D)$ 同构.

注　两个域 F 和 F' 之间的同构是指, 把 F 和 F' 看作交换环时它们之间的同构. 特别地, 它是 F 和 F' 之间满足下列性质的一一对应, 即如果 $x \leftrightarrow x'$ 和 $y \leftrightarrow y'$, 那么

$$(x+y) \leftrightarrow (x'+y') \quad \text{和} \quad (xy) \leftrightarrow (x'y').$$

证明　域 F 包含商 $\frac{a}{b}$, 这个商是方程 $bx = a$ 的解, 其系数 a 和 $b \neq 0$ 在 D 中, 所有这些商的集合 S 包含所有整数 $\frac{a}{1} = a$. 根据定理 2 中的法则, S 在加法、减法、乘法和除法之下是封闭的, 于是, 在 F 的这些运算之下, S 可以描述成 D 的闭包. 总之, S 是一个域 (定理 3).

这些商 $\frac{a}{b}$ 以定理 2 的 (i)~(iii) 所描述的方式进行相加、相乘以及表示相等, 完全相同的法则用到数偶 (a, b) 上, 因此对应 $\frac{a}{b} \leftrightarrow (a, b)$ 是 D 的闭包 S 到 $Q(D)$ 上的一个同构.　证毕

特别注意, 这个对应把 D 中每个 a 映上到 $\frac{a}{1} \leftrightarrow (a, 1) = a$.

联合定理 6 和前面的推论我们得到

定理 7　*整数环 $\mathbf{Z}$ 可以按照一种且只有一种方式被嵌入域 $\mathbf{Q} = Q(\mathbf{Z})$ 中, 使得 $\mathbf{Q}$ 的每个元素是两个整数的商.*

这就完成了由整数环 $\mathbf{Z}$ 构造有理数域 $\mathbf{Q}$.

习　题

1. 详细证明: 数偶乘法的交换律和结合律.
2. 证明: 由 (5) 所定义的 "相等" 关系满足自反律、对称律和传递律.
3. 设 $\mathbf{Z}[i]$ 是所有复数 $a + bi$ 的集合, 这里 a 和 b 为整数, $i^2 = -1$.
   (a) 清楚地叙述两个这样的数是怎样相加和相乘的.
   (b) 证明它们构成一个整环.　　(c) 描述它的商域.
4. 模 6 整数环 $\mathbf{Z}_6$ 能够嵌入一个域吗? 为什么?
5. 描述模 5 整数环 $\mathbf{Z}_5$ 的商域.
6. 域 $\mathbf{Q}$ 的商域是什么? 把它一般化.
7. 证明: 在两个域之间的任何同构 $F \leftrightarrow F'$ 之下, 由 $a \leftrightarrow a'$, $b \leftrightarrow b'$ 和 $c \leftrightarrow c'$(假定 $c \neq 0$) 可推出 $c^{-1} \leftrightarrow c'^{-1}$ 和 $\frac{a-b}{c} \leftrightarrow \frac{a'-b'}{c'}$(见 1.12 节习题 1).
8. 证明: 对应 $a + b\sqrt{7} \leftrightarrow a + b\sqrt{11}$($a, b$ 为有理数) 不是同构.

*9. 证明: 形为 $a + b\sqrt{7}$ 的数构成的域 $\mathbf{Q}(\sqrt{7})$ 与形为 $a + b\sqrt{11}$ 的数构成的域 $\mathbf{Q}(\sqrt{11})$ 之间不存在同构 (a, b 为有理数). (提示: 证明没有元素能与 $\sqrt{7}$ 对应.)

10. 关于从同构的整环 D 和 D' 产生的 $\frac{a}{b}$ 和 $\frac{a'}{b'}$ 所构成的商域, 你能说些什么? 并证明你的命题.

*11. 证明：既不是 0 也不是 ±1 的任何有理数可以唯一地表示成 $\pm p_1^{e_1}\cdots p_r^{e_r}$ 的形式，其中 p_i 是正素数，适合 $p_1 < p_2 < p_3 < \cdots < p_r$，指数 e_i 是正整数或负整数.

*12. 证明：任何有理数 $\frac{r}{s} \neq 0$ 可以唯一地表示成

$$\frac{r}{s} = b_1 + \frac{b_2}{2!} + \frac{b_3}{3!} + \cdots + \frac{b_n}{n!}$$

的形式，其中 n 为适当的整数，每个 b_k 是整数，适合 $0 \leqslant b_k < k$，若 $k > 1$，并且 $b_n \neq 0$.

13. 设 p 为固定的素数，证明：适合 n 与 p 互素的所有有理数 $\frac{m}{n}$ 的集合 $\mathbf{Z}_{(p)}$ 是一个整数. $\mathbf{Z}_{(p)}$ 与它的商域是一致的.

14. 找出 $\mathbf{Q}$ 中包含有理数 $\frac{1}{6}$ 和 $\frac{1}{5}$ 的最小子整环.

*15. 描述 $\mathbf{Q}$ 的所有可能的子整环.

16. 证明：任何恰有两个元素的域与 $\mathbf{Z}_2$ 同构.

17. 设整环 $\mathbf{Z}[\sqrt{3}]$ 是由所有数 $a + b\sqrt{3}$ 组成，其中 a 和 b 为整数，证明这个整环有一个商域，它同构于所有形为 $r + s\sqrt{3}$ 的实数组成的集合，其中 r 和 s 是有理数，并得到一个明显的同构.

2.3 联立线性方程

一个域不一定由通常的“数”组成，比如，如果 p 为素数，则所有模 p 的整数就构成一个只包含有限多个不同 (即不同余) 元素的域. 整环 $\mathbf{Z}_p$ 是域这个事实是下面定理的推论.

定理 8 任何有限整环 D 是一个域.

证明 D 是有限的这个假设意味着 D 的元素全部可以列出来，排成 $b_1, b_2, \cdots, b_n$，此外 n 为某正整数 (一般有限集的讨论见第 12 章). 为证明 D 是域，我们只须证明 D 的任意指定的元素 $a \neq 0$ 在 D 中有一个逆元素. 考察所有的乘积

$$ab_1, ab_2, \cdots, ab_n (b_1, b_2, \cdots, b_n\text{为}D\text{中元素}). \tag{9}$$

这给出 D 中几个全不相同的元素，因为不然，如果对某 $i \neq j$ 有 $ab_i = ab_j$，则根据消去律，消去 a 留下 $b_i = b_j$，这与假定 $b_i(i = 1, 2, \cdots, n)$ 是不同元素相违背. 因为 D 中的全部元素都列在表 (9) 中，D 的单位元素 1 也必出现在表中某个位置上，比如 $1 = ab_i$，那么相应的元素 b_i 就是所要求的元素 a 的逆元素. 证毕

根据上述证明，为在 $\mathbf{Z}_p$ 中精确地找出逆元素，可以对 $\mathbf{Z}_p$ 中所有可能的数 b_i 进行试验来得到. 逆元素还可以直接算出，这是因为 $\mathbf{Z}_p$ 中方程 $ax = 1$(其中 $a \neq 0$) 只不过是同余方程 $ax \equiv 1 (\mathrm{mod}\ p)$(其中 $a \neq 0$) 的另一种形式，而且后者可以根据 1.9 节定理 16 所述的欧几里得算法求出整数解 x.

值得注意：联立线性方程的整个理论应用到一般域上. 例如考虑两个联立方程

$$\begin{aligned} ax + by &= e, \\ cx + dy &= f, \end{aligned} \tag{10}$$

其中字母 $a, \cdots, f$ 表示域 F 的任意元素. 第一个方程乘以 d, 第二个方程乘以 b, 然后相减, 我们得到 $(ad-bc)x = de - bf$; 第二个方程乘以 a, 第一个方程乘以 c, 然后相减, 得到 $(ad-bc)y = af - ce$, 因此, 如果我们定义 (10) 的系数行列式 (参见第 10 章) 为

$$\Delta = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc,$$

当 $\Delta \neq 0$ 时, 则方程 (10) 有解

$$x = \frac{de - bf}{\Delta}, \quad y = \frac{af - ce}{\Delta} \quad (\Delta = ad - bc), \tag{10'}$$

而且没有其他解. 但是当 $\Delta = 0$ 时, 方程 (10) 或者没有解或者有很多解 (后者仅当 $c = ka, d = kb, f = ke$ 时才发生, 也就是两个方程是"成比例"的).

高斯 (Gauss) 消去法　前面消去法的方法可以推广到形为

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{11}$$

的 n 个未知数 $x_1, \cdots, x_n$ 的 m 个联立线性方程, 这里不仅已知系数 a_{ij}, b_i, 而且未知数 x_i 都被限制在指定的域 F 上. 为求出已知方程组的全部解, 我们现在将叙述称为高斯消去法的一般方法, 其想法是用简单的方程组代替已知方程组, 这个简单的方程组等价于已知方程组, 即它们是同解方程组. (例如, 退化方程 $0 \cdot x_1 + \cdots + 0 \cdot x_n = b_i$与$0 = b_i$ 等价, 而 $0 = b_i$ 是不一定能满足的.)

采用缩写记号, 我们只写下第 i 个方程, 并把它表示成样本项 $a_{ij}x_j$对$j = 1, \cdots, n$ 求和, 即写成

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i = 1, \cdots, m; \quad \text{所有} \quad a_{ij} \in F. \tag{11'}$$

我们分两种情况对未知数的个数 n 用归纳法进行论证.

情况 1　每个 $a_{i1} = 0$. 那么显然方程组 (11′) 等价于 $n-1$ 个未知数 $x_2, \cdots, x_n$ 的 m 个方程的一个"较小的"方程组; 对于较小的方程组的任何解来说, x_1 是任意的.

情况 2 某一个 $a_{i1} \neq 0$. 通过两个方程的调换 (如果必要的话), 我们得到等价的方程组, 使得 $a_{11} \neq 0$. 当第一个方程乘以 a_{11}^{-1} 时, 我们则得到一个等价的方程组, 其中 $a_{11} = 1$. 然后, 依次从第 i 个方程 $(i = 2, \cdots, m)$ 减去新的第一个方程的 a_{i1} 倍, 我们便得到形如

$$\begin{aligned} x_1 + a'_{12}x_2 + a'_{13}x_3 + \cdots + a'_{1n}x_n &= b'_1, \\ a'_{22}x_2 + a'_{23}x_3 + \cdots + a'_{2n}x_n &= b'_2, \\ &\vdots \\ a'_{m2}x_2 + a'_{m3}x_3 + \cdots + a'_{mn}x_n &= b'_m \end{aligned} \tag{12}$$

的等价方程组. 例如, 在域 $\mathbf{Z}_{11}$ 上, 方程组

$$\begin{aligned} 3x + 5y + 7z &\equiv 6, \\ 5x + 9y + 6z &\equiv 7, \\ 2x + y + 4z &\equiv 3. \end{aligned}$$

用这个方法将化为

$$\begin{aligned} x + 9y + 6z &\equiv 2, \\ 8y + 9z &\equiv 8, \\ 5y + 3z &\equiv 10. \end{aligned}$$

这里所有方程都理解成是模 11 的.

对 m 用归纳法进行论证, 我们得到

定理 9 *任意 n 个未知数 m 个方程的联立线性方程组 (11) 可化为一个等价的方程组, 这个等价方程组的第 i 个方程具有形式*

$$x_i + c_{i,i+1}x_{i+1} + c_{i,i+2}x_{i+2} + \cdots + c_{in}x_n = d_i, \tag{13}$$

这里 i 属于 $\{1, 2, \cdots, m\}$ 中 r 个数组成的某个子集, 然后再加上 $m-r$ 个形为 $0 = d_k$ 的方程.

证明 如果总是出现情况 2, 则我们得到形为 (12) 的 m 个方程, 并且称原方程组是相容的. 如果出现情况 1, 则我们可以得到形为 $0 = d_k$ 的一组退化方程. 如果所有的 $d_k = 0$, 则可以不必考虑 $0 = d_k$ 的那些方程, 如果有一个 $d_k \neq 0$, 则原方程组 (11) 是不相容的 (没有解). 证毕

详细写出方程组 (13) 如下

$$\begin{aligned} x_1 + c_{12}x_2 + c_{13}x_3 + \cdots + c_{1n}x_n &= d_1, \\ x_2 + c_{23}x_3 + \cdots + c_{2n}x_n &= d_2, \\ &\vdots \\ x_r + \cdots + c_{rn}x_n &= d_r \qquad (r \leqslant m), \end{aligned}$$

可称为梯形方程组.

任何梯形方程组 (13) 的解法是容易描述的. 逐次考虑 $x_n, x_{n-1}, \cdots, x_1$. 如果在该序列中出现的 x_i 是方程组 (13) 中某个方程的第一个变量, 那么它可通过 $x_n, \cdots, x_{i+1}$ 由下面关系确定出来

$$x_i = d_i - c_{i,i+1}x_{i+1} - c_{i,i+2}x_{i+2} - \cdots - c_{in}x_n. \tag{13'}$$

否则, 这个 x_i 可以取任意值. 这就证明了

推论　*在定理 9 所说的相容情况下, (11) 的全部解确定如下. 不出现在 (13) 各式首位的 $m-r$ 个变量 x_k 可以任意取值 (它们是自由变量). 任意选取这些 x_k 之后, 代入 (13′) 式便可逐步地算出剩下的变量 x_i.*

在前面举出的具体例子中, 首先 $8y+9z \equiv 8 \pmod{11}$ 可化为 $y + 8z \equiv 1 \pmod{11}$, 这个方程乘以 5 后去减方程 $5y + 3z \equiv 10 \pmod{11}$, 我们得 $7z \equiv 5 \pmod{11}$, 因此 $z \equiv 7 \pmod{11}$. 于是已知方程组的梯形方程组是

$$\left.\begin{aligned} x + 9y + 6z &\equiv 2 \\ y + 8z &\equiv 1 \\ z &\equiv 7 \end{aligned}\right\} \pmod{11}.$$

我们解得 $y \equiv 1 - 8z \equiv 0 \pmod{11}$ 和 $x \equiv 2 - 9y - 6z \equiv 4 \pmod{11}$. 将 $x = 4, y = 0, z = 7$ 代入原方程, 可以验证它是原方程的解.

如果 (11) 右边的常数 b_i 全都为零, 则称方程组为*齐次的*. 这类方程组总有 (平凡) 解 $x_1 = x_2 = \cdots = x_n = 0$. 它可能不存在非平凡解, 但是如果变量的个数超过方程的个数, 那么方程组 (13) 的最后一个方程总还包含可任意取值的自由变量. 此外, 对于齐次方程组来说, 决不会出现可以矛盾的方程 $0 = d_i$, 因此有

定理 10　*n 个变量 m 个方程的齐次线性方程组, 当 $m < n$ 时, 总有非全为零的解.*

习　题

1. 解下列联立同余式:

(a) $\begin{cases} 3x + 2y \equiv 1 \\ 4x + 6y \equiv 3 \end{cases} \pmod 7$;

(b) $\begin{cases} 2x + 7y \equiv 3 \\ 3x + 4z \equiv 6 \\ 4x + 7y + z \equiv 0 \end{cases} \pmod{11}$;　(c) $\begin{cases} x - 2y + z \equiv 5 \\ 2x + 2y \equiv 7 \\ 5x - 3y + 4z \equiv 1 \end{cases} \pmod{13}$.

2. 删去习题 1(a) 和 (b) 方程中的模, 在有理数域 $\mathbf{Q}$ 中求解.

3. 在 $\mathbf{Q}(\sqrt{2})$ 中解联立方程

$$\begin{cases}(1+\sqrt{2})x+(1-\sqrt{2})y=2,\\(2-\sqrt{2})x+(3-\sqrt{2})y=1.\end{cases}$$

4. 求出下列联立同余式的全部非同余解：

$$\begin{cases}x+y+z\equiv 0\\3x+2y+4z\equiv 0\end{cases}\pmod 5.$$

5. 求出下列联立同余式的全部非同余解：

$$\text{(a)}\begin{cases}x+2y-z+5t\equiv 4\\2x+5y+z+2t\equiv 1\\x+3y+2z+6t\equiv 2\end{cases}\pmod 7;\qquad \text{(b)}\begin{cases}x+y+z\equiv 1\\3x+3y+3z\equiv 4\end{cases}\pmod 5.$$

6. 证明：两个方程

$$a_1x_1+\cdots+a_nx_n=c,\qquad b_1x_1+\cdots+b_nx_n=d$$

总有解, 这里系数在给定的域中, 并假定不存在常数 $k\neq 0$ 和 $m\neq 0$ 使得对于 $i=1,\cdots,n$, 有 $ka_i=mb_i$.

7. 证明：如果 $(x_1,\cdots,x_n)$ 是齐次线性方程组的任意解, 那么 $(-x_1,\cdots,-x_n)$ 是另一个解. 关于这两个解的和能说些什么?

*8. (a) 证明：三个联立方程

$$ax+by+cz=d,\quad a'x+b'y+c'z=d',\quad a''x+b''y+c''z=d''$$

在任意域 F 上有唯一解, 这里 3×3 行列式

$$\Delta=ab'c''+a'b''c+a''bc'-a''b'c-a'bc''-ab''c'\neq 0.$$

(b) 写出 (a) 中计算 x 的公式, 并用它证明 $x=4$ 是 (12) 式下面列出的 $\mathbf{Z}_{11}$ 上三个联立线性方程的解.

2.4 有 序 域

如果域 F 包含 "正" 元素集合 P, 1.3 节中所列的加法律、乘法律和三分律成立, 则称域 F 是有序的. 换句话说, 当把一个域看作整环时, 如果它是一个有序整环, 则这个域是有序域. 根据经验知道, 全体有理数就构成这样的有序域, 现在我们从构造有理数为整数偶出发来证明这一点, 并进一步指出, 这种 "自然" 排序的方法, 是把有理数域作成有序域的唯一方法.

首先回忆一下, 任何有序整环中, 非零元素 b 的平方 b^2 总是正的. 如果商 $\frac{a}{b}$ 是正的, 则乘积 $\left(\frac{a}{b}\right)b^2 = ab$ 也必为正的, 反之亦真. 因此, 在任意有序域中,

$$\frac{a}{b} > 0 \quad \text{当且仅当} \quad ab > 0, \tag{14}$$

而有理数 (a,b) 的意思是表示商 $\frac{a}{b}$, 因此我们定义有理数 (a,b) 是正的当且仅当在 $\mathbf{Z}$ 中乘积 ab 是正的.

定理 11 *如果定义 $(a,b)>0$ 意味着整数 ab 是正的, 则全体有理数构成一个有序域.*

证明 我们按前面的习惯定义了相等之后, 必须证明与正元素相等的元素是正的: 由 $(a,b)>0$ 和 $(a,b)\equiv(c,d)$ 推出 $(c,d)>0$. 这是正确的, 因为 cd 与 b^2cd 同号, ab 与 abd^2 同号, 根据假设 $ad=bc$, 有 $abd^2=b^2cd$. 所需的加法律、乘法律和三分律也成立. 例如, 两个正的数偶 (a,b) 与 (c,d) 的和是正的, 这是因为, 由 $ab>0$ 和 $cd>0$ 推出 $d^2ab>0$和$b^2cd>0$, 因此

$$bd(ad+bc) = d^2ab + b^2cd > 0,$$

这就是说, 和 $(ad+bc,bd)$ 是正的. 最后, 分数 "正" 元素的定义同表示整数的特殊分数 $(a,1)$ 的自然顺序是一致的, 这是因为, 根据定义 (14), 只有当 $a\cdot 1>0$ 时, $(a,1)$ 才是正的. 证毕

因为定理 11 的证明中只用到 "全体整数是有序整环" 的假定, 所以它实际上建立了下面更一般的结果:

定理 12 *在约定 "D 的元素 a,b 的商是正的当且仅当 ab 是正的" 之下, 有序整环 D 的商域 $Q(D)$ 是有序的. 只有按这种方法可以扩张 D 的次序使 Q 成为有序域.*

存在很多其他有序域: 实数域、形为 $a+b\sqrt{2}$ 的域 $\mathbf{Q}(\sqrt{2})$(见 2.1 节) 和实数域的其他子域, 在任何这样的域中, 绝对值可按 1.3 节那样定义, 在那里所建立的不等式的性质在这里同样成立. 在任何有序域上, 除任意有序整环上成立的法则之外, 我们还可以证明,

$$0 < \frac{1}{a} \text{ 当且仅当 } a > 0, \tag{15}$$

$$\frac{a}{b} < \frac{c}{d} \text{ 当且仅当 } abd^2 < b^2cd, \tag{16}$$

$$\text{由 } 0 < a < b \text{ 可推出 } 0 < \frac{1}{b} < \frac{1}{a}, \tag{17}$$

$$\text{由 } a < b < 0 \text{ 可推出 } 0 > \frac{1}{a} > \frac{1}{b}, \tag{18}$$

$$a_1^2 + a_2^2 + \cdots + a_n^2 \geqslant 0. \tag{19}$$

(17) 和 (18) 两个法则在不等式除法中是常见的. 法则 (19), 即平方和永远非负 (1.3 节定理 2), 是特别有用的. 例如, 若 $a \neq b$, 则 $(a-b)^2 > 0$, 于是 $a^2 - 2ab + b^2 > 0$, 由此得出 $a^2 + b^2 > 2ab$. 由此令 $x = a^2, y = b^2$, 并且两边除以 2, 那么

$$\frac{x+y}{2} > \sqrt{xy} \quad (x \neq y).$$

这表明, 两个不同实数的算术平均值大于几何平均值 $\sqrt{xy}$.

习　题

1. 假定全体整数构成一个有序整环, 证明两个正有理数的乘积是正的.
2. 类似地证明: 设 D 是有序整环, 如果 $(a,b) \neq 0$, 那么 $(a,b) > 0$ 和 $-(a,b) > 0$ 两种情况中, 恰有一种情况在 $Q(D)$ 中成立.
3. 证明

$$|xx' + yy'| \leqslant \sqrt{(x^2+y^2)(x'^2+y'^2)}$$

在任何有序域中成立, 该域中所有正元素都有平方根. (提示: 两边平方.)

4. 证明正文中的公式 (15) 至公式 (19).
5. 设 n 为正整数, a 和 b 为正有理数, 证明 $\frac{a^n+b^n}{2} \geqslant \left(\frac{a+b}{2}\right)^n$.(提示: 令 $\frac{a+b}{2} = r, d = \frac{a-b}{2}, a = r+d, b = r-d$.)
6. (a) 证明: 任何有序域的子域是有序域.

   (b) 有序域的任何子整环是有序整环吗?
7. 对全体有理数 (或者更一般地, 在任何有序域中) 证明: 如果 $a < b$. 则存在无穷多个 x 满足 $a < x < b$.
8. 证明: 在序域中, 正元素不能构成一个良序集.
9. 在算术中常常发生的错误是把 $\frac{a}{b} + \frac{a}{c}$ 计算成 $\frac{a}{b+c}$.

   (a) 证明: 在任何域中, 由 $\frac{a}{b} + \frac{a}{c} = \frac{a}{b+c}$ 可推出 $a = 0$ 或者 $b^2 + bc + c^2 = 0$.

   (b) 证明: 在一个有序域中, 由它可推出 $a = 0$.

*2.5　正整数公设[①]

虽然我们用了全体整数的整环 **Z** 作为我们考察基本数系的出发点, 但是这一过程实际上很不严格, 因为它假定负数存在. 本章余下部分我们将指出怎样仅由我们熟悉的正整数的事实导出负整数及其性质, 由此我们指出, 负数存在性的假定如何可以避免.

① 加 “*” 号的节可以略去, 这并不影响连贯性.

为一致起见，我们从列举所有正整数系 $\mathbf{Z}^+$ 的一些基本性质开始，这些性质容易从第 1 章的结果推出.

定理 13 $\mathbf{Z}$ 中所有正整数系 $\mathbf{Z}^+$，具有下列性质：

(i) 在所定义的加法和乘法二元运算之下，$\mathbf{Z}^+$ 是封闭的，这两个运算满足结合律、交换律和分配律.

(ii) 在 $\mathbf{Z}^+$ 中存在乘法单位元素 1，适合对 $\mathbf{Z}^+$ 中一切 m 有 $m \cdot 1 = m$.

(iii) 此外，在 $\mathbf{Z}^+$ 中，下面的消去律成立：

$$\text{若}\quad mx = nx, \quad \text{则}\quad m = n. \tag{20}$$

(iv) 再有，对 $\mathbf{Z}^+$ 中任意两个元素 m 和 n，下面三种可能性中恰有一个成立：或者 $m = n$，或者 $m + x = n$ 在 $\mathbf{Z}^+$ 中有一个解 x，或者 $m = n + y$ 在 $\mathbf{Z}^+$ 中有一个解 y.

(v) 最后，在 $\mathbf{Z}^+$ 中数学归纳法原理成立：$\mathbf{Z}^+$ 的任意子集，如果包含 1，并且当它包含 n 时也包含 $n+1$，那么这个子集包含 $\mathbf{Z}^+$ 中每个元素.

我们把 $\mathbf{Z}^+$ 的这些性质的证明留作习题.

相反地，如果把这个定理中指出的性质 (i)~性质 (v) 看作公设，在下述意义下，它们完整地描述了正整数：我们先前定义过的正整数系具有这些性质，并且可以证明任何其他满足这些公设的系统与这个正整数系同构．特别注意，在 $\mathbf{Z}^+$ 中如果 $m + x = n$，那么

$$n + z = (m + x) + z = m + (x + z) = m + (z + x) = (m + z) + x,$$

因此，由 (iv) 知 $m + z = n + z$ 是不可能的. 类似地，$n = m + y$ 同 $m + z = n + z$ 也是不相容的. 因此，第三次利用性质 (iv)，我们得到

$$\text{若}\quad m + z = n + z, \quad \text{则}\quad m = n. \tag{21}$$

而且，方程 $m + x = n$ 的三种可能性代替了正整数系的那些序的性质.

从由这些公设给出的正整数系出发，我们可以重新构造整数系统 $\mathbf{Z}$. 构造的目的是为了得到一个比 $\mathbf{Z}^+$ 大的系统，在这个系统中减法总是可能的. 因此，作为新元素，我们引进某正整数偶 (m, n)，这里每个数偶表示方程 $n + x = m$ 的解 (如果是解的话). 这个构造的详细过程类似于由整数环构造有理数域 (2.2 节).

定义 一个整数定义为正整数 m 和 n 的一个数偶 (m, n). 数偶的“相等”按约定定义为

$$(m, n) \equiv (r, s) \quad \text{意味着}\quad m + s = n + r, \tag{22}$$

而和与积分别定义为

$$(m, n) + (r, s) = (m + r, n + s), \tag{23}$$

$$(m, n) \cdot (r, s) = (mr + ns, ms + nr). \tag{24}$$

最后, (m, n) 是"正"的当且仅当对某正整数 x 有 $n + x = m$.

由这些定义引进的数偶实际上满足我们已给出的所有关于整数的公设. 我们首先必须验证, 由 (22) 引进的"相等"满足自反律、对称律和传递律. 在这个相等意义下, 分别由 (23) 和 (24) 给出的和与积是唯一确定的. 把定义 (23) 和 (24) 系统地应用到整环的各种形式的定律上, 那么这些定律对于数偶也成立, 这同有理数的讨论几乎一样. 特别地, 对刚刚定义的系统, (2, 1) 是单位元素, (1, 1) 是零元素, 并且加法逆元素存在, 这是因为

$$(m, n) + (n, m) \equiv (1, 1), \quad \text{对所有} \quad (m, n).$$

数偶的乘法消去律证起来较难些, 证明时用到定理 13 的条件 (iv). 证明完消去律之后, 我们就知道全体数偶构成整环.

由定理 13 的公设 (iv), 每个数偶恰好可写成下面三种形式之一: $(m, m), (n + x, n), (m, m + x)$. 第一种形式的那些数偶等于零元素 (1, 1); 第二种形式的数偶 $(n + x, n)$ 是正的数偶. 并且可以证明, 数偶具有有序整环的定义 (1.3 节) 中所要求的加法律、乘法律和三分律. 此外,

$$(m + x, m) \equiv (n + y, n) \quad \text{当且仅当} \quad x = y.$$

因此, 如果把"同余"的数偶实际上看成同一偶, 那么对应 $x \longmapsto (n + x, n)$ 是从全体给定的正整数 x 的集合到全体新的正数偶 $(n + x, n)$ 的集合的一个单射. 它甚至是一个单一同态, 这因为由定义 (23) 和 (24),

$$(m + x, m) + (n + y, n) = (m + n + x + y, m + n),$$
$$(m + x, m) \cdot (n + y, n) = (mn + my + nx + mn + xy, mn + nx + mn + my).$$

因此新的"正"数偶满足数学归纳法原理. 于是我们就粗略地给出了下面结果的一个证明.

定理 14 通过定义 $\mathbf{Z}$ 的任意元素为 $\mathbf{Z}^+$ 中两个正整数之差这种方式, 正整数系 $\mathbf{Z}^+$ 可以嵌入较大的系统 $\mathbf{Z}$ 中, 在 $\mathbf{Z}$ 中减去是可能的. 这样构造的系统 $\mathbf{Z}$ 是一个有序整环, 它的正元素满足数学归纳法原理.

由 1.5 节习题 8, 这个结果蕴含着良序原则. 值得注意的是, 上面粗略的证明只涉及 $\mathbf{Z}^+$ 的公设, 反过来, 在包含 $\mathbf{Z}^+$ 的任意整环中, $\mathbf{Z}^+$ 的元素之差 $(a - b)$ 必须满足定义 (22)~(24)(参见 1.2 节习题 5). 这就证明了

定理 15 包含系统 $\mathbf{Z}^+$ 的任意整环包含一个与整数环 $\mathbf{Z}$ 同构的子整环.

习 题

1. 证明：由 (22) 定义的关系满足自反律、对称律和传递律.
2. 证明：如果 $(m,n) \equiv (m',n')$, 则对所有的 (r,s) 有

$$(m,n)+(r,s) \equiv (m',n')+(r,s), \quad 和 \quad (m,n)\cdot(r,s) \equiv (m',n')\cdot(r,s).$$

3. 证明：由 (23) 定义的"加法"满足交换律和结合律.
4. 证明：由 (24) 定义的"乘法"满足交换律和结合律.
5. 证明：对所有的 $m, (m,m)$ 是同一个元素, 并且是加法零元素. 证明第一个命题可由第二个命题推出.
6. 证明：$(m+1,m)$ 是乘法单位元素.
7. 证明分配律.
8. 证明乘法消去律.
9. 习题 1~ 习题 8 中用到 $\mathbf{Z}^+$ 哪些性质? 像由定理 5 得出定理 7 那样, 叙述一个与定理 14 和定理 15 有关的定理.
10. 证明：对于数偶 (m,n) "正性"的任何不同于 (24) 式后面一段叙述的定义, 定理 14 都不成立.
11. 详细证明定理 13.

*12. 证明：定理 13 的公设 (iv), 可以用条件"对 $\mathbf{Z}^+$ 中每个 $m, m+1 \neq 1$"来代替.(这实质上是皮亚诺 (Peano) 公设 (iii), 如定理 16 所述.)

13. 在 $\mathbf{Z}^+$ 中定义 $m<n$ 意味着对某个 $x \in \mathbf{Z}^+$ 有 $m+x=n$. 证明：
    (a) 由 $m<n$ 和 $n<r$ 可推出 $m<r$.
    (b) 不存在 m 使 $m<m$ 成立.
    (c) 由 $m<n$ 可推出对所有的 $r, m+r<n+r$.
    (d) 由 $m<n$ 可推出对所有的 $r, mr<nr$.

*14. 证明：习题 13 的结论 (c) 和 (d) 可以用来代替 $\mathbf{Z}^+$ 的公设中的消去律 (20) 和 (21).

15. 证明：由 $\mathbf{Z}^+$ 得到 $\mathbf{Z}$ 所用的方法并不能产生 $\mathbf{Z}$ 的新扩张, 你能推广这个结果吗?

*2.6 皮亚诺公设

在正整数集合 $P=\mathbf{Z}^+$ 上, 如果把加法和乘法当作未定义的运算, 我们可以用后继函数

$$S(n)=n+1 \tag{25}$$

来定义它们.

定理 16 正整数集合 P 和后继函数 S 具有下列性质：

(i) $1 \in P$;

(ii) 若 $n \in P$, 则 $S(n) \in P$;

(iii) 在 P 中没有一个 n 使 $S(n) = 1$;

(iv) 对 P 中 m 和 n, 由 $S(n) = S(m)$ 可推出 $n = m$;

(v) P 的一个子集如果包含 1, 并且当它包含 n 时, 也包含 $S(n)$, 那么这个子集必等于 P.

证明 这些性质直接从定理 13 得到. 特别注意, (v) 是数学归纳法原理. 证毕

性质 (i)~性质 (v) 称为正整数集合的*皮亚诺公设*. 正如下面将要指出的那样, 它们足以证明正整数集的所有性质. 我们现在用它们证明, 原来的整数公设可确定整数集合 (精确到同构).

定理 17 *在任意有序整环 D 中, 存在唯一的子集 P' 满足关于单位元素 $1'$ 和后继函数 $S'(a) = a + 1'$ 的皮亚诺公设.*

注 直观地, 显然由 $2' = 1' + 1', 3' = 1' + 1' + 1', \cdots$ 定义的序列 $1', 2', 3', \cdots$ 就是这样一个子集. 不过, 我们希望一个以有序整环公设为依据的正式的证明.

证明 D 的所有正元素的集合 D^+ 显然包含 $1'$, 并且满足 (i) 和 (ii). 现在令 Σ 是 D^+ 的所有子集 T 组成的类, 而 T 具有 P 中性质 (i) 和 (ii), 我们定义 P' 是所有这些集合 T 的交集, 即 $a \in P'$ 当且仅当 a 属于每一个这样的集合 T.

由定义, 对于 P', (i) 和 (ii) 成立. 因为 P' 只包含正元素, 所以 (iii) 成立; 因为 $a + 1' = b + 1'$ 意味着 $a = b$, 所以 (iv) 成立. 为证明 (v), 令 A 是 P' 的子集, 它包含 $1'$, 并且当它包含 a 时也包含 $S'(a)$. 那么 A 是前面用到的集合 T 中的一个, 于是 P' 包含在 A 中, 因此 $P' = A$. 对 P', 这就证明了 (v), 同时 (v) 表明 P' 是唯一可能的这样的集合, 因为 P' 满足 (i) 和 (ii).

定理 18 *定理 17 的子集 P' 对于加法、乘法和序而言, 它同构于正整数集合 P.*

注 非正式地, 显然 $1 \longmapsto 1', 2 \longmapsto 2', \cdots$ 产生所要求的同构. 因为 $1' < 1' + 1' < 1' + 1' + 1' < \cdots$, 所以这个对应将保持次序.

证明 首先, 令 $Q(n)$ 是如下命题: P 中整数 $1 \leqslant x \leqslant n$ 和 P' 中元素 $\phi_n(x)$ 之间存在唯一的对应 $x \longmapsto \phi_n(x)$, 在这对应下:

$$\phi_n(1) = 1', \quad \phi_n(S(x)) = S'(\phi_n(x)), \text{ 对 } \quad 1 \leqslant x < n. \tag{26}$$

显然 $Q(1)$ 成立. 已知 $Q(n)$ 成立, 因此有一个 ϕ_n, 我们可以通过令

$$\phi_{n+1}(x) = \phi_n(x), \text{ 对 } 1 \leqslant x \leqslant n \text{ 和 } \phi_{n+1}(n+1) = S'(\phi_n(n)),$$

构造唯一的 ϕ_{n+1}, 因此由 $Q(n)$ 成立推出 $Q(n+1)$ 成立. 由归纳法这就证明了 $Q(n)$ 成立.

再有, 如果 $1 \leqslant x \leqslant n < m$, 对 x 用归纳法, 我们可以证明 $\phi_n(x) = \phi_m(x)$, 因此当 $x \leqslant n$ 时,$\phi_n(x)$ 是不依赖于 n 的. 令 $\phi(x)$ 表示 P' 的元素, 这就给出 P 到 P' 的

对应 $x \longmapsto \phi(x)$, 它具有性质:

$$\phi(1) = 1', \qquad \phi(S(x)) = S'(\phi(x)). \tag{27}$$

P' 的每个元素是 P 的某个元素 x 的对应元素 $\phi(x)$. 因为元素 $\phi(x)$ 的集合包含 $1'$, 并且包含任何 $\phi(x)$ 也一定包含 $\phi(x)$ 的后继, 因此根据 P' 的性质 (v), 这个集合就是整个 P'.

在两个集合 P 和 P' 中, 我们有

$$n + 1 = S(n), \quad n + S(m) = S(n + m), \tag{28}$$

$$n \cdot 1 = n, \qquad n \cdot S(m) = n \cdot m + n. \tag{29}$$

从这些方程和 (27) 式, 对 m 用归纳法, 我们可以容易地证明

$$\phi(n + m) = \phi(n) + \phi(m) \quad 和 \quad \phi(n \cdot m) = \phi(n) \cdot \phi(m).$$

换句话说, ϕ 关于加法和乘法是一个同构.

其次, ϕ 保留次序, 即由 $m < n$ 可推出 $\phi(m) < \phi(n)$. 实际上, 由定义, $m < n$ 意味着 $n - m$ 是正的, 即

$$m < n \quad 当且仅当 \quad n = m + k, \quad 对P中某个k. \tag{30}$$

因此由 $m < n$ 得出 $n = m + k$, 所以 $\phi(n) = \phi(m) + \phi(k)$. 因为 $\phi(k)$ 在 D 中是正的, 所以这就证明了 $\phi(m) < \phi(n)$, 正如所要求的那样.

最后, ϕ 是 P 到 P' 的双射. 因为我们已经知道, $\phi(x)$ 的集合包含整个 P', 所以只须证明, 由 $n \neq m$ 可推出 $\phi(n) \neq \phi(m)$. 但是,$n \neq m$ 的意思是, 比如说是 $m < n$, 于是 $\phi(m) < \phi(n)$, 因此

$$\phi(n) \neq \phi(m).$$

为概括我们的结论, 我们定义两个有序整环之间的*序–同构*, 即保留次序的同构. 鉴于定理 15, 我们从定理 18 得出下列推论:

推论 1　*任意有序整环包含一个与* $\mathbf{Z}$ *序–同构的子整环.*

这个结果同定理 6 和定理 7 结合起来, 我们有

推论 2　*任意有序域包含一个与有理数域* $\mathbf{Q}$ *序–同构的子域.*

这个结果给出作为最小有序域的有理数域的一个抽象特征.

最后, 在 D 中全体正元素集合是良序的情况下, 可以容易地证明, 定理 17 的集合 P' 是由 D 的所有正元素组成. 这就证明了:

推论 3　*在序–同构意义下, 只存在一个有序整环* $\mathbf{Z}$, *它的正元素构成良序集合.*

这就证明了，在同构意义下，整数公设唯一地确定整数集合.

可以不从良序整环公设开始论述整数集合，而是从皮亚诺公设开始. 其要点是注意到可用递归方程 (28) 和 (29) 定义完备的加法表和乘法表. 同定理 15 的证明中几乎一样，我们可以正式地证明存在唯一的满足 (28) 的二元运算 —— 加法，类似地，存在唯一的满足 (29) 的乘法. 那么定理 13 中列举的各种性质可用归纳法证明. 然后，从皮亚诺公设出发，2.5 节中给出的数偶构造产生全体整数.

习　题

在下列习题中，只假定皮亚诺公设，并由方程 (28) 和方程 (29) 定义了加法和乘法.

1. 用归纳法证明 $n+1=1+n$.
2. 利用习题 1 证明，加法满足交换律.
3. 证明：加法满足结合律.
4. 证明：乘法满足结合律.
5. 证明：分配律成立.

第3章 多 项 式

3.1 多项式形式

设 D 为任意整环, 设 x 是较大的整环 E 的任意元素, D 作为 E 的子整环包含在 E 中. 在 E 中, 我们能作 x 同 D 的元素或同 x 本身的和、差与积.

反复进行这些运算, 明显得到下面形式的一切表达式

$$a_0 + a_1x + \cdots + a_nx^n \quad (a_0, \cdots, a_n \in D; a_n \neq 0, \text{ 当 } n > 0), \tag{1}$$

这里 x^n(n 为任意整数) 定义为 n 个因子的乘积 $xx\cdots x$. 而反过来, 只用整环公设, 我们可对形为 (1) 的任意两个表达式进行加、减与乘, 得到第三个这样的表达式. 例如, 如果 D 是整数环, 则根据一般分配律、交换律和结合律, 有

$$\begin{aligned} f(x) &= (0 + 1 \cdot x + (-2)x^2)(2 + 3 \cdot x) \\ &= 0 \cdot 2 + 0 \cdot 3 \cdot x + 1 \cdot x \cdot 2 + 1 \cdot x \cdot 3 \cdot x \\ &\quad +(-2)x^2 \cdot 2 + (-2)x^2 \cdot 3 \cdot x \\ &= 0 + 0 \cdot x + 2x + 3x^2 + (-4)x^2 + (-6)x^3 \\ &= 0 + (0+2)x + (3 + (-4))x^2 + (-6)x^3 \\ &= 0 + 2x + (-1)x^2 + (-6)x^3. \end{aligned}$$

这个论证可以一般化. 事实上, 设

$$\begin{aligned} & p(x) = a_0 + a_1x + \cdots + a_mx^m \\ \text{和} \quad & q(x) = b_0 + b_1x + \cdots + b_nx^n \end{aligned}$$

是形为 (1) 的任意两个表达式. 如果 $m > n$, 那么我们有

$$p(x) \pm q(x) = (a_0 \pm b_0) + \cdots + (a_n \pm b_n)x^n + a_{n+1}x^{n+1} + \cdots + a_mx^m. \tag{2}$$

如果 $m \leqslant n$, 可得出类似的公式. 再有, 根据分配律,

$$p(x)q(x) = \sum_{i=0}^{m}\sum_{j=0}^{n} a_ib_jx^{i+j}.$$

然后把指数相同的项集中在一起, 并将系数相加, 我们有

$$p(x)q(x) = a_0b_0 + (a_0b_1 + a_1b_0)x + \cdots + a_mb_nx^{m+n}. \tag{3}$$

在这个公式中, x^k 的系数显然是和

$$\sum_i a_ib_{k-i},$$

这里是对满足 $0 \leqslant i \leqslant m$ 和 $0 \leqslant k-i \leqslant n$ 的所有 i 求和, 见图 3-1.

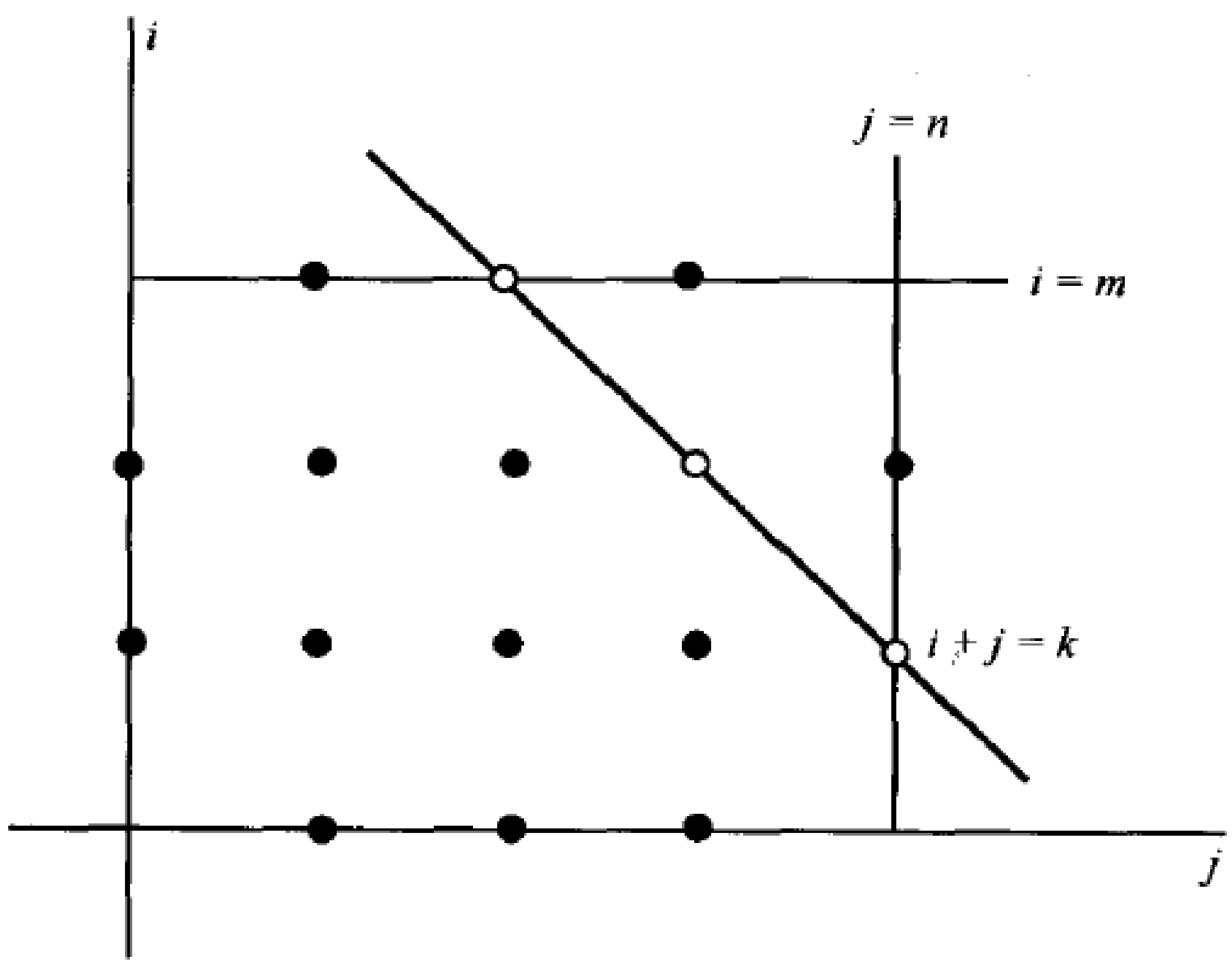


图 3-1

于是我们就证明了下面的结果:

定理 1 *假设存在一个整环 E, 包含一个与给定整环 D 同构的子整环, 并有元素 x 不在 D 中. 那么关于这个元素 x 的多项式 (1) 根据公式 (2) 和公式 (3) 相加、相减和相乘, 构成 E 的子整环.*

为了证明这样的整环 E 总是存在的, 需要建立下面的定义.

定义 *整环 D 上关于 x 的多项式是指形为 (1) 的表达式. 整数 n 称为多项式 (1) 的次数. 两个多项式相等是指它们具有相同的次数, 而且对应的系数都相等.*

因为关于符号 x 没有给出什么假定, 所以表达式 (1) 也常称为*多项式形式*(这里把它同多项式函数加以区别, 见 3.2 节), 符号 x 本身称为*未定元*.

定理 2 *如果加法和乘法分别由公式 (2) 和公式 (3) 定义, 那么整环 D 上关于 x 的全体不同的多项式形式构成一个包含 D 在内的新整环 $D[x]$.*

证明 由公式 (3) 推出没有零因子 (乘法消去律), 这是因为, 两个非零多项式形式乘积的首项系数 a_mb_n 是它相应因子的非零首项系数 a_m 和 b_n 的乘积 (非零的).0 和 1 的性质及加法逆元素的存在性不难从公式 (2) 和公式 (3) 得出.

为了证明交换律、结合律和分配律, 引进 "哑" 零系数是方便的, 这使 (2) 和 (3) 变成简单形式

$$\sum_{k=0}^{\infty} a_k x^k + \sum_{k=0}^{\infty} b_k x^k = \sum_{k=0}^{\infty} (a_k + b_k) x^k, \tag{2'}$$

$$\left(\sum_{k=0}^{\infty} a_k x^k\right)\left(\sum_{k=0}^{\infty} b_k x^k\right) = \sum_{k=0}^{\infty}\left(\sum_{i+j=k} a_i b_j x^k\right), \tag{3'}$$

这里除了有限多个系数外全都是零. 那么任何一个定律, 比如分配律, 只要把定律的两边按照法则 (2′) 和 (3′) 乘起来就可以验证, 因为,

$$\left(\sum_k a_k x^k\right)\left(\sum_k b_k x^k + \sum_k c_k x^k\right) = \sum_k \left[\sum_{i+j=k} a_i (b_j + c_j)\right] x^k,$$

$$\left(\sum_k a_k x^k\right)\left(\sum_k b_k x^k\right) + \left(\sum_k a_k x^k\right)\left(\sum_k c_k x^k\right)$$

$$= \sum_k \left[\left(\sum_{i+j=k} a_i b_j\right) + \left(\sum_{i+j=k} a_i c_j\right)\right] x^k,$$

并证明这两个等式右边的 x 的每个幂 x^k 的系数相等. 根据整环 D 的分配律, 两个表达式中 x 的 k 次幂的系数是相同的. 类似的论证可证其余定律, 从而完成定理 2 的其他证明.

现在回忆一下 2.2 节的定理 7, 我们会看到, 如果我们定义 D 上关于未定元 x 的*有理形式*为带有非零分母多项式形式的形式商

$$\frac{p(x)}{q(x)} = \frac{a_0 + a_1 x + \cdots + a_m x^m}{b_0 + b_1 x + \cdots + b_n x^n} \quad (a_i, b_j \text{ 在 } D\text{中};\ a_m \neq 0,\ \text{当 } m > 0; b_n \neq 0),$$

并由 2.2 节的定义 (5),(6) 和 (7) 来分别定义有理形式的相等、加法和乘法, 这样我们便可得到一个域.

推论 *任意整环 D 上关于未定元 x 的有理形式构成一个域. 这个域记作 $D(x)$.*

习 题

1. 把下列各式化成形式 (1): $x^2 - 5x(3x+7)^2$,

   $(x^2+5x-4)(x^2-2x+3)$, $\left(3x^2+7x-\frac{1}{2}\right)\left(x^3-\frac{x}{2}+1\right)$.

2. 类似习题 1, 计算 $(3x^3+5x-4)(4x^3-x+3)$, 其中系数是 mod 7 的整数.
3. x^3+5x-4 是 (1) 的形式吗? 把它化成形式 (1). 把

$$(1+x+2x^2+3x^3)-(0+x+x^2+3x^3)$$

化成形式 (1), 指出每一步用什么公设.

4. (a) $\frac{1}{2}+3\cdot x^{\frac{1}{2}}+5x$ 是有理数域上的多项式形式吗?

   (b) 在系数属于 $\mathbf{Z}_5$ 的多项式形式整环上, 为什么 $x^3\cdot x^4$ 不等于 x^2?
5. 讨论下列命题:

   (a) 两个多项式形式乘积的次数等于这两个因子的次数之和.

   (b) 两个多项式形式之和的次数等于被加数的次数较大者.
6. 证明: 在 $D[x]$ 中, 加法结合律和乘法结合律成立.
7. $p(x)=a_0+a_1x+\cdots+a_nx^n$ 的 "形式导数" 定义为 $p'(x)=a_1+2a_2x+\cdots+na_nx^{n-1}$. 证明: 在任意整环上:

   (a) $(cp)'=cp'$(c为常数), (b) $(p+q)'=p'+q'$,

   (c) $(pq)'=pq'+p'q$, (d) $(p^n)'=np^{n-1}p'$.

*8. 设 $p(y)$ 和 $q(x)$ 分别是关于未定元 y 和 x 的多项式形式, 证明: 把 $y=q(x)$ 代入 $p(y)$, 产生一个多项式 $p(q(x))$. 根据习题 7 中定义的形式导数, 证明: $[p(q(x))]'=p'(q(x))\cdot q'(x)$.

*9. 对于给定的整环 D 指出, 怎样构造由关于符号 t 的全体 "形式" 的无穷幂级数 $a_0+a_1t+a_2t^2+\cdots$(其中系数 a_i 在 D 中) 组成的整环 $D\{t\}$.

*10. (a) 设 D 为一个有序整环, 证明: 如果规定多项式形式 $p(x)>0$ 是指 $p(x)$ 的第一个非零系数 a_k 是正的 (在 D 中), 那么多项式形式 (1) 构成有序整环 $D[x]$.

   (b) 证明: 如果我们规定 $p(x)>0$ 是指在形式 (1) 中 $a_n>0$, 那么 $D[x]$ 也是有序整环.

*11. 在习题 10(b) 中, 令 $D=\mathbf{Z}$, 证明: 1 是 $\mathbf{Z}[x]$ 中最小 "正" 多项式, 虽然 $\mathbf{Z}[x]$ 并不满足良序原则.

3.2 多项式函数

如前所述, 设 D 为任意整环, 又设

$$f(x)=a_0+a_1x+\cdots+a_mx^m$$

是 D 上关于 x 的任意多项式形式. 如果未定元 x 用一个元素 $c\in D$ 代替, $f(x)$ 就不再是一个虚的表达式, 它可以看作 D 中一个确定元素

$$a_0+a_1c+\cdots+a_mc^m.$$

换句话说, 如果 x 被看作在微积分学的意义下的一个独立变量, 而不是看作 D 外面的抽象符号, 那么 $f(x)$ 就成为普通的函数:"如果 x 已知 (值为 c), 那么 $f(x)$ 就被确定了 (值为 $f(c)$)". 我们把它抽象化, 一般地定义变量在 D 上的"函数" f 是一个规则: 它给 D 上每个元素 x 确定一个"值" $f(x)$, 这个值也在 D 中. 我们定义两个这样的函数相等 (记作 $f = g$) 当且仅当对所有的 $x, f(x) = g(x)$. 两个函数的和$h = f + g$, 差 $q = f - g$ 及积 $p = fg$ 分别通过对所有的 x 计算 $h(x) = f(x) + g(x), q(x) = f(x) - g(x)$和$p(x) = f(x)g(x)$ 来定义的. 常值函数是取值 b 与 x 无关的函数; 恒等函数 是函数 j, 它满足对所有 $x, j(x) = x$.

定义 多项式函数是可以写成形式 (1) 的函数.

因为推导公式 (2) 和 (3) 时所用的法则在任何整环中都是成立的, 所以不管未定元 x 取什么值①c(在 D 中), 公式 (2) 和 (3) 都成立. 也就是说, 它们是恒等式, 因此多项式函数的和与积也可以通过公式 (2) 和 (3) 来计算. 正如 3.3 节将要说明的那样, 按 1.1 节的定义, D 上全体多项式函数构成一个交换环.

根据定义, 每个形式 (1) 都确定一个唯一的多项式函数, 每个多项式函数至少由一个这样的形式来确定. 因此无疑存在一个保持和与积的映射, 它把任意给定整环 D 上的全体多项式形式的集合映射到全体多项式函数的集合.(这样的对应称为映上同态或满同态, 见 3.3 节.)

如果可以确定映射是一一的, 我们就知道它是一个同构. 因此, 从抽象代数的观点来看, 我们可以忽略多项式形式与多项式函数之间的差别. 可惜情况并非如此. 事实上, 在模 3 整数域 $\mathbf{Z}_3$ 上, $f(x) = x^3 - x$ 和 $g(x) = 0$ 这两个不同的形式确定了同一个函数 —— 这个函数恒等于零. 根据费马定理 (1.9 节定理 18), 在 $\mathbf{Z}_p$ 上, $x^p - x$ 与 0 是相同的. 因此, 在任意 $\mathbf{Z}_p$ 上, 多项式函数相等实际上不同于多项式形式相等.

我们现在将指出, 在上述例子中, 由于系数所在的整环是有限的, 发生这一事实, 并不奇怪. 在有理数域上, 我们并不能构造出一个这样的例子. 我们在说明此事之前先回忆一些基本定义. 所谓非零形式 (1) 的次数, 我们指的是它的最大指数, 即 n. 最高次项 $a_n x^n$ 称为它的首项, a_n 称为它的首项系数, 如果 $a_n = 1$, 多项式则称为首一多项式.

定理 3 整环 D 上的一个多项式形式 $r(x)$ 可被 $x - a$ 整除当且仅当 $r(a) = 0$.

这里"$r(x)$ 可被 $x - a$ 整除"这句话的意思是 $r(x) = (x - a) \cdot s(x)$, 其中 $s(x)$ 是 D 上的某一个多项式形式.

证明 设 $r(x) = c_0 + c_1 x + \cdots + c_n x^n (c_n \neq 0)$. 对每个 a, 由中学代数公式, 我们

① 实际上, 通过设"x 为未知量"解方程的根据是: 在 x 上所允许的每个运算, 对于每个可能的 x 值都必须是正确的.

有

$$\sum_{k=0}^{n} c_k x^k - \sum_{k=0}^{n} c_k a^k = \sum_{k=0}^{n} c_k (x^k - a^k)$$
$$= \sum_{k=0}^{n} c_k [(x-a)(x^{k-1} + x^{k-2}a + \cdots + a^{k-1})].$$

因此 $r(x) - r(a) = (x-a)s(x)$, 其中 $s(x)$ 是 $n-1$ 次多项式形式. 反之, 如果 $r(x) = (x-a)s(x)$ 中用 a 代替 x, 则得 $r(a) = 0$.

推论 *整环 D 上的 n 次多项式 $r(x)$ 在 D 中至多有 n 个零点.*

($r(x)$ 的零点是指方程 $r(x) = 0$ 的根, 即元素 $a \in D$ 使得 $r(a) = 0$.)

证明 如果 a 是一个零点, 那么根据定理有 $r(x) = (x-a) \cdot s(x)$, 其中 $s(x)$ 的次数为 $n-1$. 由归纳法, $s(x)$ 至多有 $n-1$ 个零点, 可是根据 1.2 节定理 1 知, $r(x) = 0$ 当且仅当 $x = a$ 或 $s(x) = 0$, 因此 $r(x) = 0$ 至多有 n 个零点.

定理 4 *如果整环 D 是无限的, 那么 D 上定义同一个函数的两个多项式形式具有相等的系数.*

证明 像 (1) 那样, 设 $p(x)$ 和 $q(x)$ 是两个给定的关于未定元 x 的多项式形式. 如果它们确定同一个函数, 那么对于 D 中选取的每个元素 a 都有 $p(a) = q(a)$; 然而我们所希望的结论则是 $p(x)$ 和 $q(x)$ 的次数相等, 对应的系数相同. 如果用差 $r(x) = p(x) - q(x)$ 来表示, 这就是说, 对 D 中一切 $a, r(a) = c_0 + c_1 a + \cdots + c_n a^n = 0$ 可推出 $c_0 = c_1 = \cdots = c_n = 0$. 这个结论可由定理 3 推出, 因为如果系数 c_i 不全为零, 那么, 在 D 中至多有 n 个 x 使多项式 $r(x)$ 为零, 因为 D 是无限的, 所以还剩下一些 x 值使 $r(x) \neq 0$, 这与 $r(x)$ 在 D 上为零相矛盾.

于是, 如果 D 是无限的, 则多项式函数和多项式形式这两个概念是等价的 (用代数学的术语就是, 多项式函数环同构于多项式形式环).

另一方面, 如果 D 是包含元素 $a_1, a_2, \cdots, a_n$ 的有限整环, 则定理 4 一定不成立. 例如, 在这种情况下, n 次首一多项式形式 $(x-a_1)(x-a_2)\cdots(x-a_n)$ 同形式 0 确定了同一个函数.

因为同构于整环的任意系统本身也是一个整环, 所以定理 4 蕴含着下面的推论:

推论 *任意无限整环上全体多项式函数构成一个整环.*

如果 D 为无限域, 则不同的有理形式确定不同的有理函数, 所以 D 上全体有理函数构成一个域. (留心, 一个有理函数不是在一切点上都有定义, 只是在那些使分母不为零的点上才有定义. 因此, 如果 D 是一个域, 那么在除有限个点外的全部点上它是有定义的.)

我们常常希望找出一个次数最小的多项式 $p(x)$, 使它在域 F 中的 $n+1$ 个已知

点 $a_0, a_1, \cdots, a_n$ 上分别取 F 中给定的值 $y_0, y_1, \cdots, y_n$, 即

$$p(a_i) = y_i \quad (i = 0, 1, \cdots, n; a_i \neq a_j, \text{ 当} i \neq j). \tag{4}$$

这称为多项式插值问题.

为了解决这个问题, 考虑多项式

$$q_i(x) = \prod_{j \neq i}(x - a_j) = (x - a_0)\cdots(x - a_{i-1})(x - a_{i+1})\cdots(x - a_n).$$

显然若 $j \neq i, q_i(a_j) = 0$, 而

$$C_i = q_i(a_i) = \prod_{j \neq i}(a_i - a_j) \neq 0.$$

因此 C_i^{-1} 存在, 并且下面这个次数为 n 或低于 n 的多项式

$$p(x) = \sum_{i=0}^{n} C_i^{-1} y_i q_i(x) = \sum_{i=0}^{n} \frac{y_i \prod_{j \neq i}(x - a_j)}{\prod_{j \neq i}(a_i - a_j)} \tag{5}$$

满足方程 (4). 公式 (5) 称为拉格朗日 (Lagrange) 插值公式.

由定理 3 知, 至多有一个 n 次或低于 n 次的多项式能够满足方程 (4): 因为两个这样的多项式之差有 $n+1$ 个零点, 于是它必为零多项式形式. 这就证明了下面的结果:

定理 5 *存在一个而且只存在一个 n 次或低于 n 次的多项式形式, 在 $n+1$ 个不同点上取给定的值.*

习 题

1. 在整环 $\mathbf{Z}_5$ 上找出另一个多项式形式同 $x^2 - x + 1$ 确定同一个函数.
2. 证明: $x^2 - 1$ 在 $\mathbf{Z}_{15}$ 上有四个零点. 为什么这与定理 3 的推论不矛盾?
3. 证明: 如果 $a_0 = a_1 - h, a_2 = a_1 + h, 1 + 1 \neq 0$, 那么 (4) 式当 $n = 2$ 时的解可由抛物线插值公式

$$p(x) = y_1 + \frac{1}{2}(y_2 - y_1)\left(\frac{x - a_1}{h}\right) + \frac{1}{2}(y_2 - 2y_1 + y_0)\left(\frac{x - a_1}{h}\right)^2$$

给出.
4. 用以下方法求满足 $f(0) = 0, f(1) = 1, f(2) = 0, f(3) = 1$ 的三次多项式 $f(x) = a + bx + cx^2 + dx^3$. 把 a, b, c, d 当作四个方程的未知数, 其中最后一个方程是 $a+3b+9c+27d = 1$. (这就是待定系数法.)

5. 用插值公式 (5) 证明：任何有限域 (如 $\mathbf{Z}_p$) 上的每个函数等于某个多项式函数.

*6. 设 D 是具有 n 个元素 $a_1, a_2, \cdots, a_n$ 的有限整环, 又设 $m(x)$ 表示固定的多项式形式 $(x-a_1)\cdots(x-a_n)$.

   (a) 证明：如果两个多项式形式 $f(x)$ 和 $g(x)$ 确定同一个函数, 那么 $m(x)$ 是形式 $f(x)-g(x)$ 的因子.

   (b) 对整环 $\mathbf{Z}_3$ 和 $\mathbf{Z}_5$ 求出 $m(x)$.

   (c) 证明：在 $D=\mathbf{Z}_p$ 的情况下, $m(x)=x^p-x$.(提示：用费马定理.)

7. 证明：在一个无限域上, 确定同一个函数的不同的有理形式在 2.2 节的意义下, 它们形式上是相等的.

8. (a) 设 D 和 D' 是同构的整环, 证明 $D[x]$ 与 $D'[y]$ 同构, 这里 $D[x]$ 和 $D'[y]$ 分别是 D 和 D' 上关于未定元 x 和 y 的多项式形式的整环.

   (b) 关于 $D(x)$ 和 $D'(y)$ 有什么结论?

9. 设 Q 是整环 D 的商域 (2.2 节定理 4), 证明：域 $D(x)$ 与域 $Q(x)$ 同构.

3.3 交换环的同态

设 D 为任意给定的整环, 又设 $D\langle x\rangle$ 表示 D 上的多项式函数集合. 对所有 $x\in D, f(x)+g(x)=g(x)+f(x), 0+f(x)=f(x), 1\cdot f(x)=f(x)$, 等等. 因此, 加法和乘法满足交换律、结合律和分配律; 加法和乘法的单位元素存在; 并且加法逆元素存在. 概括起来, $D\langle x\rangle$ 除乘法消去律外满足整环的所有公设. 当 D 为有限整环时, 消去律不成立, 这因为存在非零因子的乘积 $(x-a_1)\cdot(x-a_2)\cdots(x-a_n)$ 为零.

换句话说, 在 1.1 节的定义下, $D\langle x\rangle$ 是一个交换环. 为方便起见, 我们在此重述这个定义.

定义 *交换环在称为加法和乘法的两种二元运算之下封闭的集合, 这两种运算满足交换律和结合律, 并且进一步有:*

(i) 满足乘法对加法的分配律;

(ii) 存在加法单位元素 (零)0, 并且存在加法逆元素;

*(iii) 存在乘法单位元素 1.*①

回忆一下, 我们已经证明了 1.2 节所列的法则 1～ 法则 9 对任意交换环都成立. 另外, 1.10 节的定理 19 构造了一类有趣的有限交换环 $\mathbf{Z}_m$.

任意整环 D 上的全体函数构成的系统 D^*, 为我们提供了另一个交换环的例子, 这里加法和乘法像 3.2 节里那样定义. 甚至于定义在无限整环 D 上的全体函数集合 D^* 中也存在零因子. 例如, 如果 D 为任意有序整环, 我们定义 $f(x)=|x|+x, g(x)=|x|-x$, 那么 $f\cdot g=h, h(x)=|x|^2-x^2=0$, 对一切 x 成立. 但是

① 一些作者在定义交换环时去掉条件 (iii). 非交换环将在第 13 章讨论.

$f \neq 0, g \neq 0$. 另一方面, D^* 具有整环定义中所有其他性质. 我们只要在每步证明的右边简单写上"对一切 x", 根据 D 的定律便可得到 D^* 的相应定律的证明. 例如, $f(x)+g(x)=g(x)+f(x)$ 对一切 x, 这就意味着 $f+g=g+f$. 再有, 如果我们定义 e 为一个常值函数, 即对一切 x, $e(x)=1$, 那么对一切 x 和 f, $e(x)f(x)=1\cdot f(x)=f(x)$, 因此对一切 f, $ef=f$, 于是 e 是 D^* 的乘法单位元素. (想一想乘法消去律为什么不能按这种方法证明.) 因为上面没有一处用到乘法消去律, 所以我们可以断言:

引理 1 任意交换环 A 上的全体函数构成一个交换环.

现在让我们定义交换环 A 的子环(类似于子整环) 为 A 的这样的子集: 如果它包含任意两个元素 f 和 g, 则它包含 $f \pm g$ 和 fg, 并且还包含着 A 的单位元素.

由定理 1, 任意整环 D 上多项式函数集合 $D\langle x\rangle$, (i) 它是 D 上所有函数环 D^* 的子环, (ii) 它包含所有常值函数和恒等函数, (iii) 它包含在任何其他满足 (ii) 的 D^* 的子环之中. 按这种意义 $D\langle x\rangle$ 是由常值函数和恒等函数生成的 D^* 的子环. 这给出多项式函数概念的一个简单的代数特征.

下面将同构的概念一般化, 可以更深刻地认识交换环.

定义 一个函数 $\phi: a \longmapsto a\phi$, 它把交换环 R 映射到交换环 R', ϕ 称为同态当且仅当它满足下列条件: 对所有 $a, b \in R$, 有

$$(a+b)\phi = a\phi + b\phi, \tag{6}$$

$$(ab)\phi = (a\phi)(b\phi), \tag{7}$$

并且把 R 的单位元素映射到 R' 的单位元素.

这些条件表明, 同态保持加法和乘法. 它们是按照 1.11 节和 1.12 节中简洁的记号写出来, 其中 $a\phi$ 表示用 ϕ 变换 a. 如果我们写 $\phi(a)$ 代替 $a\phi$, 则 (6) 和 (7) 分别变成 $\phi(a+b)=\phi(a)+\phi(b)$ 和 $\phi(ab)=\phi(a)\phi(b)$. 显然, 一个同构恰是一个双射的同态.

我们容易验证, 从 n 到包含 n 的剩余类 (对任意固定模 m) 的函数是一个同态 $\mathbf{Z} \to \mathbf{Z}_m$, 它把整数环 $\mathbf{Z}$ 映上到 1.10 节定理 19 的环 $\mathbf{Z}_m$. 我们现在证明另一个容易的结果.

引理 2 设 ϕ 是从交换环 R 到交换环 R' 的同态, 那么 0ϕ 是 R' 的零元素, 并且对所有 $a, b \in R$ 有 $(a-b)\phi = a\phi - b\phi$.

证明 由 (6) 式, $0\phi = (0+0)\phi = 0\phi + 0\phi$, 这就证明了 0ϕ 是 R' 中的零元素. 类似地, 如果 $x = a - b$ 在 R 中, 那么 $b + x = a$, 并且 $a\phi = (b+x)\phi = b\phi + x\phi$, 于是 $x\phi = a\phi - b\phi$ 在 R' 中.

定理 6 从任意整环 D 上的多项式形式整环 $D[x]$ 到 D 上多项式函数环 $D\langle x\rangle$ 的对应 $p(x) \longmapsto f(x)$ 是一个同态.

证明 对 D 中任意元素 x, D 中的元素 $p(x)$ 和 $q(x)$ 的加法和乘法必须遵循恒等式 (2) 和 (3), 因为在 3.1 节中这些恒等式的推导只用到整环公设.

定理 4 的结果指出, 如果 D 是无限的, 那么定理 6 的同态就是一个同构.

习 题

1. (a) 证明: 在域 $\mathbf{Z}_2$ 上只存在四个不同的函数, 并写出这个函数环的加法表和乘法表.
   (b) 把这些函数中的每一个表示成多项式函数.
   (c) 这个函数环与模 4 整数环同构吗?
2. 在模 n 整数环 $\mathbf{Z}_n$ 上有多少不同的函数?
3. 下列函数集合是含有单位元素的交换环吗?
   (a) 整环 D 上满足 $f(0)=0$ 的所有函数 f.
   (b) 整环 D 上满足 $f(0)=f(1)$ 的所有函数 f.
   (c) 整环 D 上满足 $f(0)\neq 0$ 的所有函数 f.
   (d) $\mathbf{Q}$ 上 ($\mathbf{Q}$ 为有理数域) 满足 $-7\leqslant f(x)\leqslant 7$(对一切 x) 的所有函数 f.
   (e) $\mathbf{Q}$ 上满足 $f(x+1)=f(x)$(对一切 x) 的所有函数 f(这样的函数是周期的).
4. 在习题 3 的那些例子之外, 再构造两个函数交换环.
5. 设 D^* 如前文中所定义的, 证明: D^* 中的加法与乘法的结合律成立.
6. (a) 设 D 和 D' 是同构的整环, 证明: $D\langle x\rangle$ 和 $D'\langle x\rangle$ 也是同构的.
   (b) 对于 D^* 和 D'^* 有什么结论?
7. 证明: 我们不能把 $\mathbf{Z}_p$ 上所有多项式函数的环 $\mathbf{Z}_p\langle x\rangle$ 嵌入一个域中.
8. 证明: 如果同态 ϕ 把交换环 R 映上到交换环 R', 那么 R 的单位元素通过 ϕ 映射到 R' 的单位元素.
9. 证明: 如果 $\phi: R\longmapsto R'$ 是环的任意同态, 那么 R 中那些映成 R' 的零元素的元素构成的集合 K 是 R 的一个子环.

*3.4 多元多项式

3.1 节 ~3.3 节的讨论只涉及单变量 (未定元)x 的多项式. 但是很多结果不难推广到多变量 (未定元)$x_1,\cdots,x_n$ 的情形.

定义 整环 D 上关于未定元 $x_1,\cdots,x_n$ 的多项式形式可递推地定义为整环 $D[x_1,\cdots,x_{n-1}]$ 上关于变量 x_n 的多项式形式, 而 $D[x_1,\cdots,x_{n-1}]$ 是 D 上关于变量 $x_1,\cdots,x_{n-1}$ 的多项式形式组成的整环 (简单地说, $D[x_1,\cdots,x_n]=D[x_1,\cdots,x_{n-1}][x_n]$). 整环 D 上关于变量 $x_1,\cdots,x_n$ 的多项式函数是由常值函数 $f(x_1,\cdots,x_n)=c$

和 n 个恒等函数 $f_i(x_1, \cdots, x_n) = x_i (i = 1, \cdots, n)$ 通过加法、减法和乘法构造出来的.

例如, 在两个变量 x, y 的情况下,

$$p(x,y) = (3 + x^2) + 0 \cdot y + (2x - x^3)y^2$$

是一个这样的形式 —— 通常把它写成更顺的形式

$$3 + x^2 + 2xy^2 - x^3y^2.$$

根据定理 4 对 n 用归纳法得

定理 7 *如果 D 是无限的, 那么 D 上关于变量 $x_1, \cdots, x_n$ 的每个多项式函数可按一种且只有一种方法表示为多项式形式. 不管 D 是无限的还是有限的, $D[x_1, \cdots, x_n]$ 是一个整环.*

从定义明显看出, 变量下标的每个置换, 引导出关于 $D\langle x_1, \cdots, x_n \rangle$ 的一个自然的自同构, 其中 $D\langle x_1, \cdots, x_n \rangle$ 是 n 个变量多项式函数的交换环. 如果 D 是无限的, 由定理 7 得到, 上述结论对多项式形式也是正确的 (这些定义对于变量不是对称的). 现在我们证明, 这个结论对任意整环 D 都是正确的.

定理 8 *变量下标的每个置换, 引导出 $D[x_1, \cdots, x_n]$ 上的不同自同构.*

证明 考虑两个未定元 x, y 的情况. $D[y, x]$ 的每个形式

$$p(y,x) = \sum_i \left(\sum_j a_{ij} y^j \right) x^i,$$

可以根据 $D[y, x]$ 中的分配律、交换律和结合律重新排列得出一个形为

$$p(y,x) = \sum_j \left(\sum_i a_{ij} x^i \right) y^j$$

的表达式. 根据这个表达式的形式, 似乎可以把它解释为整环 $D[x, y]$(先 x 后 y) 中的多项式 $p'(x, y)$. 这样建立的对应 $p(y, x) \longmapsto p'(x, y)$ 是一一的 —— 每个非零元素 a_{ij} 的有限集合恰好对应 $D[y, x]$ 中一个元素, 也恰好对应 $D[x, y]$ 中一个元素. 最后, 因加法和乘法的法则 (2) 和 (3) 可以从整环的公设推出, 而 $D[y, x]$ 和 $D[x, y]$ 都是整环, 所以我们看到这个对应保持了和与积.

n 个未定元的情况可以用更复杂更一般的记号类似地处理, 或者从两个变量的情况出发用归纳法推导出.

于是 $D[x_1, \cdots, x_n]$ 事实上对称地依赖于 $x_1, \cdots, x_n$. 这就启发我们构造一种 $D[x_1, \cdots, x_n]$ 的定义, 从这定义对称性是一目了然的. 在 $n = 2$ 情况下对于整环 $D'' = D[x, y]$, 可以粗略地说明如下. 第一, D'' 是由 x, y 和 D 的元素生成的 (D'' 的

每个元素可以由 x, y 和 D 的元素反复进行求和与求积运算而得). 第二, 生成元 x 和 y 是 D 上并立未定元 (或者是在 D 上代数独立的). 这就意味着, 系数 a_{ij} 在 D 上的有限和 $\sum_{i,j} a_{ij}x^iy^j$ 可以为零当且仅当所有系数 a_{ij} 全为零. 这两个性质以对称的方式 (见下面的习题 9) 唯一确定整环 $D[x, y]$.

习　题

1. 把下列各式表示成系数在 $D[x]$ 上关于 y 的多项式:
   (a) $p(x, y) = y^3x + (x^2 - xy)^2$,
   (b) $q(x, y) = (x + y)^3 - 3yx(x^2 + x - 1)$.
2. 计算整环 $\mathbf{Z}_2$ 上关于两个变量 x, y 的所有可能的函数的数目.
3. 重写下列表达式为关于 x 的多项式, 其系数是关于 y 的多项式 (像定理 8 的证明中那样):
$$(3x^2 + 2x + 1)y^3 + (x^4 + 2)y^2 + (2x - 3)y + x^4 - 3x^2 + 2x.$$
4. 设 D 为任意整环, 证明: 把 $p(x)$ 映射到 $p(-x)$ 的对应是 $D[x]$ 的一个自同构. 它也是 $D\langle x\rangle$ 的自同构吗?
5. 对应 $p(x) \longmapsto p(x + c)$(此处 c 为常数) 是 $D[x]$ 的自同构吗? 用数的例子加以说明.
6. 设 F 为域, 证明: 对任意常数 $a \neq 0$, 对应 $p(x) \longmapsto p(ax)$ 是 $F[x]$ 的一个自同构.
7. 除了定理 8 中所叙述的外, 列出 $D[x, y]$ 上的自同构.
8. 证明定理 7:
   (a) 对 $n = 2$,　　(b) 对任意 n.
9. (a) 详细证明: 整环 $D[x, y]$(先 x 后 y) 确实是由两个并立未定元 x 和 y 在 D 上生成的.
   (b) 设 D' 和 D'' 是两个整环, 它们分别是由两个并立未定元 x', y' 和 x'', y'' 在 D 上生成的, 证明: 在 "x' 映射到 x'', y' 映射到 y'', D 的每个元素映射到它自身" 的对应之下, D' 和 D'' 同构.
   (c) 对 $n = 2$, 利用 (a) 和 (b) 给出定理 8 一个另外的证明.

3.5 辗转相除法

多项式辗转相除法 (有时称为 "多项式长除法") 为下面多项式相除提供了一个标准形式: 用一个多项式 $a(x)$ 去除另一个多项式 $b(x)$ 以便得到商式 $q(x)$ 和余式 $r(x)$, $r(x)$ 的次数低于除式 $a(x)$ 的次数. 我们现在将证明, 这个辗转相除法, 虽然通常是在有理系数多项式上进行的, 但实际上对于系数在任意域上的多项式都是可行的.

定理 9 如果 F 为任意域, $a(x) \neq 0$ 和 $b(x)$ 是 F 上的任意多项式, 那么我们可以找到 F 上的多项式 $q(x)$ 和 $r(x)$, 使得

$$b(x) = q(x)a(x) + r(x) \tag{8}$$

成立, 这里 $r(x)$ 或者为零或者它的次数低于 $a(x)$ 的次数.

证明概要 从 $b(x)$ 中减去除式 $a(x)$ 与适当的单项式 cx^k 的乘积, 逐步消去被除式 $b(x)$ 的最高项. 如果 $a(x) = a_0 + a_1x + \cdots + a_mx^m (a_m \neq 0), b(x) = b_0 + b_1x + \cdots + b_nx^n (b_n \neq 0)$, 并且 $b(x)$ 的次数 n 不低于 $a(x)$ 的次数 m, 则我们可以做差

$$\begin{aligned} b_1(x) &= b(x) - \frac{b_n}{a_m}x^{n-m}a(x) \\ &= 0 \cdot x^n + \left(b_{n-1} - \frac{a_{m-1}b_n}{a_m}\right)x^{n-1} + \cdots, \end{aligned} \tag{9}$$

$b_1(x)$ 的次数低于 n 或者为零. 然后我们可以重复这一过程直到余式的次数低于 m 为止.

辗转相除法的正式证明可以根据数学归纳法第二原理, 如 1.5 节中所描述的那样. 设 m 是 $a(x)$ 的次数. 任何次数 $n < m$ 的多项式 $b(x)$ 可表示成 $b(x) = 0 \cdot a(x) + b(x)$, 其商式 $q(x) = 0$. 对次数 $n \geqslant m$ 的多项式, 由 (9) 式得到

$$b(x) = b_1(x) + \frac{b_n}{a_m}x^{n-m}a(x), \tag{10}$$

其中 $b_1(x)$ 的次数 $k < n$, 除非 $b_1(x) = 0$, 由数学归纳法第二原理我们可以假定, 表达式 (8) 对于一切次数 $k < n$ 的多项式都成立, 于是我们有

$$b_1(x) = q_1(x)a(x) + r(x), \tag{11}$$

其中 $r(x)$ 的次数低于 m, 除非 $r(x) = 0$. 把 (11) 式代入 (10) 式中, 我们得到所要求的方程 (8)

$$b(x) = \left[q_1(x) + \frac{b_n}{a_m}x^{n-m}\right]a(x) + r(x).$$

特别是, 如果多项式 $a(x) = x - c$ 是线性首一的, 那么 (8) 中的余式是一个常数 $r = b(x) - (x - c)q(x)$, 如果我们令 $x = c$, 这个方程给出 $r = b(c) - 0q(c) = b(c)$. 因此我们有

推论 用 $x - c$ 去除多项式 $p(x)$, 其余数是 $p(c)$ (称为余数定理).

当 (8) 中的余式是零时, 我们就称 $b(x)$ 可被 $a(x)$ 整除. 更确切地说, 如果 $a(x)$ 和 $b(x)$ 是整环 D 上的两个多项式形式, 那么, 在 D 上 (或在 $D[x]$ 中) $b(x)$ 可被 $a(x)$ 整除当且仅当存在某个多项式形式 $q(x) \in D[x]$, 使得 $b(x) = q(x)a(x)$.

习 题

1. 证明：在 (8) 式中对于给定的 $a(x)$ 和 $b(x)$, $q(x)$ 和 $r(x)$ 是唯一的.
2. 设 $b(x)=x^5-x^3+3x-5$, $a(x)=x^2+7$, 计算 $q(x)$ 和 $r(x)$.
3. 设 $a(x)$ 分别为 $x-2, x+2, x^3+x-1$, $b(x)$ 同习题 2 一样, 计算 $q(x)$ 和 $r(x)$.
4. (a) 对域 $\mathbf{Z}_5$ 做习题 2.　　(b) 对域 $\mathbf{Z}_3$ 做习题 3.
5. 在域 F 中给出不同的数 $a_0, a_1, \cdots, a_n$, 令 $a(x)=\prod\limits_{j=0}^{n}(x-a_j)$. 证明：$F$ 上任意多项式 $f(x)$ 被 $a(x)$ 除所得的余式 $r(x)$, 确是 $f(x)$ 在这些点上的拉格朗日插值.
6. 在 $\mathbf{Z}_3, \mathbf{Z}_5, \mathbf{Z}_7$ 中任何一个整环上 x^3+x^2+x+1 可被 x^2+3x+2 整除吗?
7. 找出所有可能的环 $\mathbf{Z}_n$, 在其上 $x^5-10x+12$ 可被 x^2+2 整除.
8. (a) 设任意整环上一个多项式 $f(x)$ 有 $f(a)=0=f(b)$, 其中 $a \neq b$, 证明：$f(x)$ 可被 $(x-a)(x-b)$ 整除.
   (b) 推广这个结果.
9. 在应用数学归纳法第二原理证明辗转相除法时,$P(n)$(见 1.5 节) 确切的含义是什么?

3.6 单位与相伴

我们可以得到关于多项式的完全类似于算术基本定理的定理 (或称为唯一因子分解定理). 在这个类比中, “不可约多项式” 扮演素数的角色,它的定义如下.

定义　*一个多项式形式如果它可以分解出系数在 F 上次数较低的多项式因子, 则称它为 F 上可约多项式; 否则称它为 F 上不可约多项式.*

例如, 多项式 x^2+4 在有理数域上是不可约的. 如果不然, $x^2+4=(x+a)(x+b)$. 令 $x=-b$ 代入上式得 $(-b)^2+4=(-b+a)(-b+b)=0$, 因此 $(-b)^2=-4$. 这显然是不可能的, 因为在这个域中一个数的平方不可能是负的. 因为在任意有序域中, 同样的论证也成立, 所以我们得出结论：在实数域或其他任意有序域上, x^2+4 也是不可约的.

为了阐明不可约多项式和素数之间的类似, 我们现在对任意整环 D 来定义某些整除性的概念, 例如对多项式环 $\mathbf{Q}[x]$、整数环 $\mathbf{Z}$ 或者别的整环来定义.

D 的元素 a 可被 b*整除*(记作 $b|a$) 的定义是, 在 D 中存在某个 c, 使 $a=cb$. 如果 $b|a$ 而且 $a|b$, 则称两个元素 a 和 b 是*相伴*. 单位元素 1 的相伴称为*单位*. 因为对一切 a, 有 $1|a$, 所以 u 是 D 中的单位当且仅当它在 D 中有乘法逆元素 u^{-1}, 使得 $1=uu^{-1}$. 具有这个性质的元素也称为*可逆元素*.

如果 a 和 b 是相伴, $a=cb$ 并且 $b=c'a$, 因此 $a=cc'a$. 由消去律得 $1=cc'$, 于是 c 和 c' 都是单位. 反过来, 如果 u 是单位, 则 $a=ub$ 是 b 的相伴. 因此两个元素是相伴当且仅当其中每一个可以从另外一个乘以单位因子而得到.

例 1 在域中, 每个 $a \neq 0$ 都是单位.

例 2 在整数环 $\mathbf{Z}$ 中, 单位只有 ± 1, 因此任何 a 的相伴是 $\pm a$.

例 3 在未定元 x 的多项式环 $D[x]$ 中, 乘积 $f(x) \cdot g(x)$ 的次数是这两个因子的次数之和. 因此任何元素 $b(x)$ 如果有多项式逆 (即 $a(x)b(x) = 1$), 它必须是零次多项式 $b(x) = b$. 这样常数多项式 b 有逆仅当 b 在 D 中有逆. 因此 $D[x]$ 的单位都是 D 的单位.

如果 F 是域, 那么多项式环 $F[x]$ 的单位恰是 F 的非零常数, 因此两个多项式 $f(x)$ 和 $g(x)$ 在 $F[x]$ 中相伴当且仅当每一个是另外一个的常数倍.

例 4 在一切数 $a+b\sqrt{2}(a, b$ 为整数) 构成的整环 $\mathbf{Z}[\sqrt{2}]$ 中, 由 $(a+b\sqrt{2})(x+y\sqrt{2}) = 1$ 得出 $x = \dfrac{a}{a^2 - 2b^2}, y = -\dfrac{b}{a^2 - 2b^2}$—— 这些都是整数当且仅当 $a^2 - 2b^2 = \pm 1$. 于是 $1 \pm \sqrt{2}$ 和 $3 \pm 2\sqrt{2}$ 是 $\mathbf{Z}[\sqrt{2}]$ 中的单位, 而 $2 + \sqrt{2}$ 不是 $\mathbf{Z}[\sqrt{2}]$ 中的单位.

任意整环 D 的元素 b 可被它的一切相伴整除, 还可被一切单位整除. 这些相伴和单位称为 b 的"假因子". 不是单位也不具有真因子的元素称为 D 中*素元素*或称它在 D 中是*不可约的*.

例 5 在任意域 F 上, 线性多项式 $ax + b(a \neq 0)$ 是不可约的, 这是因为它的因子只是常数 (单位) 或是它本身的常数倍 (相伴).

例 6 考虑"高斯整数"环 $\mathbf{Z}[\sqrt{-1}]$, 它是由所有形为 $a + b\sqrt{-1}$(其中 $a, b \in \mathbf{Z}$) 的数组成. 如果 $a + b\sqrt{-1}$ 是单位, 那么对某个 $c + d\sqrt{-1}$, 我们有

$$1 = (a + b\sqrt{-1})(c + d\sqrt{-1}) = (ac - bd) + (ad + bc)\sqrt{-1}.$$

因此 $ac - bd = 1, ad + bc = 0$, 并容易验证

$$1 = (ac - bd)^2 + (ad + bc)^2 = (a^2 + b^2)(c^2 + d^2).$$

因为 $a^2 + b^2, c^2 + d^2$ 都是非负整数, 所以我们推断 $a^2 + b^2 = c^2 + d^2 = 1$; 于是只可能是: $1, -1, \sqrt{-1}$ 和, $-\sqrt{-1}$, 给出四个单位.

引理 *在任意整环 D 中, 关系"a 和 b 是相伴"是一个等价关系.*

证明将留给读者 (还见下面的习题 1 ~ 习题 3).

习 题

1. 在任意整环 D 中, 证明:
   (a) 关系"$b|a$"满足自反律和传递律.
   (b) 如果 $c \neq 0$, 那么 $b|a$ 当且仅当 $bc|ac$.
   (c) 任意两个元素有公因子和公倍数.
   (d) 如果 $a|b$ 和 $a|c$, 那么 $a|(b \pm c)$.
2. 证明: $\mathbf{Z}_m$ 的单位都是与 m 互素的整数.

3. 在任意整环中, 设“$a \sim b$”的含意是“a 和 b 相伴”, 证明:
   (a) 如果 $a \sim b$, 那么 $c|a$ 当且仅当 $c|b$.
   (b) 如果 $a \sim b$, 那么 $a|c$ 当且仅当 $b|c$.
   (c) 如果 $a|c$ 当且仅当 $b|c$, 那么 $a \sim b$.
   (d) 如果 p 为素元素并且 $p \sim q$, 那么 q 也是素元素.
4. 证明: 如果 $a \sim a'$ 且 $b \sim b'$, 那么 $ab \sim a'b'$. 而一般来说, $a+b \sim a'+b'$ 是不对的.
5. 证明广义消去律: 如果 $ax \sim by, a \sim b$ 并且 $a \neq 0$, 那么 $x \sim y$.
6. 列出 x^2+2x-1 在 $\mathbf{Z}_5[x]$ 中的所有相伴.
7. 找出两个未定元的多项式环 $D[x,y]$ 中的全部单位.
8. 对于整环 D 中哪些元素 a, 使得对应 $p(x) \longmapsto p(ax)$ 是 $D[x]$ 的自同构?
9. 找出整环 D 中的全部单位, 这里 D 是由所有有理数 $\dfrac{m}{n}$ 组成, 其中 m 和 n 为整数, 并且 n 不能被 7 整除.
10. 当 $\alpha = a + b\sqrt{3}$, 定义 $N(\alpha) = a^2 - 3b^2$, 证明:
    (a) $N(\alpha\alpha') = N(\alpha)N(\alpha')$.
    (b) 如果 α 是 $\mathbf{Z}[\sqrt{3}]$ 中的单位, 那么 $N(\alpha) = \pm 1$.
11. 设 $\mathbf{Z}[\sqrt{5}]$ 是由一切数 $\alpha = a + b\sqrt{5}$(a, b 为整数) 组成的整环, 且令 $N(\alpha) = a^2 - 5b^2$.
    (a) 证明: $9 + 4\sqrt{5}$ 是这个整环中的单位 (参见习题 10).
    (b) 证明: $1 - \sqrt{5}$ 和 $3 + \sqrt{5}$ 是相伴, 但不是单位.
    (c) 证明: 一般地, α 是单位当且仅当 $N(\alpha) = \pm 1$.
    (d) 设 $N(\alpha)$ 是 $\mathbf{Z}$ 中的素元素, 证明: α 是 $\mathbf{Z}[\sqrt{5}]$ 中的素元素.
    (e) 证明: $4 + \sqrt{5}$ 和 $4 - \sqrt{5}$ 是素元素.
    (f) 证明: 2 和 $3 + \sqrt{5}$ 是素元素. (提示: 对任何 $x \in \mathbf{Z}, x^2 \equiv 2 (\mathrm{mod} 5)$ 是不可能的.)
    (g) 利用 $2 \cdot 2 = (3 + \sqrt{5})(3 - \sqrt{5})$ 证明: $\mathbf{Z}[\sqrt{5}]$ 不是唯一因子分解整环 (见 3.9 节).
12. 详细证明正文中的引理.

3.7 不可约多项式

多项式代数中的一个基本问题是寻求判断给定域上多项式可约性的有效方法, 这种判断自然完全依赖于所考虑的域 F. 例如, 在复数域 $\mathbf{C}$ 上, 多项式 x^2+1 分解为 $x^2+1=(x+\sqrt{-1})\cdot(x-\sqrt{-1})$. 事实上, 正如 5.3 节中将要指出的, $\mathbf{C}[x]$ 中只有线性多项式是不可约的. 而 x^2+1 在实数域 $\mathbf{R}$ 上是不可约的.

再有, 因为 $x^2-28=(x-\sqrt{28})(x+\sqrt{28})$, 所以多项式 x^2-28 在实数域上是可约的. 但是, 这个多项式在有理数域上是不可约的. 后面我们将严格证明它.

引理 *一个二次或三次多项式 $p(x)$ 在域 F 上是不可约的, 除非对某个 $c \in F$, 有 $p(c)=0$.*

证明 把 $p(x)$ 任意分解成次数较低的多项式, 其中一个因子必是线性的, 这因为多项式乘积的次数等于全体因子的次数之和.

定理 10 设 $p(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n$ 是整系数多项式. 方程 $p(x)=0$ 的任何有理根 $\frac{r}{s}$ 必满足 $r|a_n$ 和 $s|a_0$.

证明 假设对某个分数 $x=\frac{b}{c}$ 满足 $p(x)=0$. 约掉 b 和 c 的最大公因子后,

我们可以把 $\frac{b}{c}$ 表示成两个互素整数 r 和 s 的商 $\frac{r}{s}$, 把它代入 $p(x)$ 得

$$0=s^np\left(\frac{r}{s}\right)=a_0r^n+a_1r^{n-1}s+\cdots+a_ns^n, \tag{12}$$

因此

$$-a_0r^n=s(a_1r^{n-1}+a_2r^{n-2}s+\cdots+a_ns^{n-1}),$$

所以 $s|a_0r^n$. 但是 $(s,r)=1$, 因此逐次应用 1.7 节定理 10 得 $s|a_0r^{n-1},\cdots,s|a_0$. 类似地, 因为

$$-a_ns^n=r(a_0r^{n-1}+\cdots+a_{n-1}s^{n-1}),$$

所以有 $r|a_n$.

推论 整系数首一多项式的任意有理根都是整数.

现在容易证明 x^2-28 在 $\mathbf{Q}$ 上是不可约的. 根据推论, $x^2=28$ 意味着 $x=\dfrac{r}{s}$ 是一个整数. 但是, 当 $|x|\geqslant 6$ 时 $x^2-28>0$, 当 $|x|\leqslant 5$ 时 $x^2-28<0$, 因此没有一个整数可能是 $x^2-28=0$ 的根, 所以 (由引理)x^2-28 在有理数域上是不可约的.

有理数域 $\mathbf{Q}$ 上多项式的不可约性的一般判别法 (容易的) 是没有的 (特殊情形的判别见 3.10 节).

习 题

1. 检验下列方程是否有有理根：
   (a) $3x^3-7x=5$, (b) $5x^3+x^2+x=4$,
   (c) $8x^5+3x^2=17$, (d) $6x^3-3x=18$.
2. 证明：$30x^n=91$(整数 $n>1$) 没有有理根. (提示：利用算术基本定理.)
3. 对哪些有理数 $x, 3x^2-7x$ 为一整数? 找出充分必要条件.
4. 0 和 250 之间有哪些整数 a, 使得对于某个 $n>1$ 方程 $30x^n=a$ 有有理根?
5. x^2+1 在 $\mathbf{Z}_3$ 上是不可约的吗? 在 $\mathbf{Z}_5$ 上呢? 对 x^3+x+2 结果如何?
6. 找出有限域使得
   (a) x^2-2 在其上是可约的.
   (b) x^2-2 在其上是不可约的.
7. 找出域 $\mathbf{Z}_5$ 上所有二次首一不可约多项式.
8. 找出域 $\mathbf{Z}_3$ 上所有三次首一不可约多项式.
9. 证明：如果 $a_0+a_1x+a_2x^2+\cdots+a_nx^n$ 是不可约的, 那么 $a_n+a_{n-1}x+\cdots+a_0x^n$ 也是不可约的.

10. 分别在下列各个域上，把多项式 x^4-5x^2+6 分解成不可约因子之积：
(a) 有理数域上. (b) 2.1 节的域 $\mathbf{Q}(\sqrt{2})$ 上. (c) 实数域上.

*11. 证明：如果 $4ac>b^2$，那么 ax^2+bx+c 在任何有序域上是不可约的.

3.8 唯一因子分解定理

整个这一节中我们将研究整环 $F[x]$ 上的因子分解，$F[x]$ 是由域F 上关于未定元 x 的多项式形成组成. 主要结果是：因子分解 (分解成不可约 (素) 因子) 是唯一的，其证明类似于算术基本定理 (第 1 章)，实际上是形式上的重复. 这个类比包含着下面基本概念，这个概念将在第 13 章里作系统讨论.

定义 *交换环 R 的非空子集 C 称为理想是指 C 满足：由 $a\in C$ 和 $b\in C$ 可推出 $a\pm b\in C$，由 $a\in C$ 和 $r\in R$ 可推出 $ra\in C$.*

注 对任意 $a\in R$，a 的所有倍数 ra 的集合是一个理想，这因为对 $r,s\in R$ 有

$$ra\pm sa=(r\pm s)a \quad 和 \quad s(ra)=(sr)a.$$

这样的理想称为*主理想*. 我们将指出，任何 $F[x]$ 中的所有理想都是主想理.

定理 11 *在任何域 F 上，$F[x]$ 的任何理想 C，(i) 或者仅由零组成，(ii) 或者由任何次数最低的非零元素 $a(x)$ 的倍数 $q(x)a(x)$ 的集合组成.*

证明 如果 $c\neq\{0\}$，则 C 包含一个次数最低的非零多项式 $a(x)$，其次数记作 $d(a)$，C 还包含 $a(x)$ 的所有倍数 $q(x)a(x)$. 这种情况下，如果 $b(x)$ 是 C 的任一多项式，则根据定理 9，有某个 $r(x)=b(x)-q(x)a(x)$ 的次数小于 $d(a)$. 但是根据假设，C 包含 $r(x)$，由 C 的构造知 C 不包含次数小于 $d(a)$ 的非零多项式，因此 $r(x)=0$，所以 $b(x)=q(x)a(x)$，这就证明了定理.

现在设 $a(x)$ 和 $b(x)$ 是任意两个多项式，考虑以任意多项式 $s(x)$ 和 $t(x)$ 作为系数的 $a(x)$ 和 $b(x)$ 所有 “线性组合” $s(x)a(x)+t(x)b(x)$ 构成的集合 C. 这个集合 C 显然是非空的，并且包含着该集元素的任意和、差或倍数，这是因为 (用缩写记号)

$$(sa+tb)\pm(s'a+t'b)=(s\pm s')a+(t\pm t')b,$$
$$q(sa+tb)=(qs)a+(qt)b.$$

因此集合 C 是一个理想，根据定理 11，它是由某个次数最低的多项式 $d(x)$ 的倍数组成.

这个多项式 $d(x)$ 将整除 $a(x)=1\cdot a(x)+0\cdot b(x)$ 和 $b(x)=0\cdot a(x)+1\cdot b(x)$，并且可被 $a(x)$ 和 $b(x)$ 的任意公因子整除，这因为 $d(x)=s_0(x)a(x)+t_0(x)b(x)$. 我们的结论是

定理 12 在 $F[x]$ 中, 任意两个多项式 a 和 b 具有最大公因子 d 满足 (i)$d|a$ 和 $d|b$, (i′) 由 $c|a$ 和 $c|b$ 可推出 $c|d$, 并且 (ii)d 是 a 和 b 的"线性组合" $d = sa + tb$.

我们注意, 可用 1.7 节中详细描述的欧几里得算法, 由 a 和 b 明确地计算出 d. (这就是上面辗转相除法可用来明显地计算多项式的余数的原因.)

还有, 如果 d 满足 (i), (i′) 和 (ii), 那么 d 的一切相伴也满足 (i), (i′) 和 (ii). 附带一句, 由 (i) 和 (ii) 可推出 (i′).

最大公因子 $d(x)$ 除单位因子外是唯一的. 这因为, 如果 d 和 d' 都是多项式 a 和 b 的最大公因子, 那么由 (i) 和 (i′), 有 $d|d'$ 和 $d'|d$, 因此 d 和 d' 确是相伴. 反之, 如果 d 是最大公因子, 那么 d 的每个相伴也是最大公因子. 有时为方便起见, 把与 d 相伴的唯一的首一多项式说成最大公因子.

两个多项式 $a(x)$ 和 $b(x)$, 如果它们的最大公因子是单位及其相伴, 则称它们*互素*. 这就意味着多项式互素当且仅当它们的公因子只能是 F 的非零常数 (整环 $F[x]$ 的单位).

定理 13 如果 $p(x)$ 是不可约的, 则由 $p(x)|a(x)b(x)$ 可推出 $p(x)|a(x)$ 或者 $p(x)|b(x)$.

证明 因为 $p(x)$ 是不可约的, 所以 $p(x)$ 和 $a(x)$ 的最大公因子或者是 $p(x)$ 或者是单位元素 1. 在前一种情况, 有 $p(x)|a(x)$, 在后一种情况, 我们可写

$$1 = s(x)p(x) + t(x)a(x),$$

因此

$$b(x) = 1 \cdot b(x) = s(x)p(x)b(x) + t(x)[a(x)b(x)].$$

因为 $p(x)$ 整除乘积 $a(x)b(x)$, 所以 $p(x)$ 整除上式右边两项, 因此整除 $b(x)$. 正如定理所要求的那样.

定理 14 $F[x]$ 中任意非常数多项式 $a(x)$ 可表示成一个常数 c 乘以某些首一不可约多项式的乘积. 这种表示除因子出现的次序外是唯一的.

证明 首先, 这样的因子分解是可能的. 如果 $a(x)$ 是常数或不可约, 那么定理显然成立. 否则, $a(x)$ 是低次多项式的乘积 $a(x) = b(x)b'(x)$. 根据数学归纳法第二原理, 我们可以假定

$$b(x) = cp_1(x)\cdots p_m(x), \qquad b'(x) = c'p'_1(x)\cdots p'_n(x),$$

因此

$$a(x) = (cc')p_1(x)\cdots p_m(x)p'_1(x)\cdots p'_n(x),$$

其中 cc' 是一常数, $p_i(x)$ 和 $p'_j(x)$ 是首一不可约多项式.

为了证明唯一性, 假设 $a(x)$ 可能有两个这样的"素"因子分解

$$a(x) = cp_1(x)\cdots p_m(x) = c'q_1(x)\cdots q_n(x).$$

显然 $c = c'$ 是 $a(x)$ 的首项系数 (因为 $a(x)$ 的首项系数是其因子首项系数之积). 再有, 因为 $p_1(x)$ 整除 $c'q_1(x)\cdots q_n(x) = a(x)$, 所以根据定理 13 它必整除某个(非常数)因子 $q_i(x)$; 因为 $q_i(x)$ 是不可约的, 所以商式 $\dfrac{q_i(x)}{p_1(x)}$ 必为常数; 又因 $p_1(x)$ 和 $q_i(x)$ 都是首一多项式, 所以常数必为 1. 因此 $p_1(x) = q_i(x)$. 消去之后, $p_2(x)\cdots p_m(x)$ 等于 $q_k(k \neq i)$ 的乘积, 并且乘积的次数低于 $a(x)$ 的次数. 因此再根据数学归纳法第二原理, $p_j(x)(j \neq 1)$ 与 $q_k(k \neq i)$ 成对地分别相等, 这就完成了证明.

一个推论是 (参考 1.8 节最后一段), 作为 $a(x)$ 的因子而出现的每个首一不可约多项式 $p_i(x)$ 的指数 e_i 是由 $a(x)$ 唯一确定的, 并且它是使得 $[p_i(x)]^e|a(x)$ 的最大的 e.

如果像定理 14 那样, 多项式 $a(x)$ 分解成不可约因子 $p_i(x)$ 的积, 但 $p_i(x)$ 不必是首一多项式, 那么, 这些因子不再是唯一的了. 然而, 每个因子 $p_i(x)$ 被它的首项系数来除而得出唯一的首一不可约因子, 因此在 $F[x]$ 上 $p_i(x)$ 是这个不可约因子的相伴. 所以, 任意两个这样的因子分解只要重新排序, 并由适当的相伴因子代替每个因子, 就可做到彼此一致. 综上所述, $F[x]$ 中的多项式的因子分解, 除了相差次序和单位因子外 (或者说除了相差次序和用相伴因子替换外) 是唯一的.

习　题

1. 证明: 如果 ϕ 是由交换环 R 到交换环 R' 的任意一个同态, 那么 R' 的加法零元素的原像构成 R 中的一个理想.
2. (a) 在 $\mathbf{Q}$ 中求出 $x^3 - 1$ 和 $x^4 + x^3 + 2x^2 + x + 1$ 的最大公因子.
   (b) 把最大公因子表示成已知多项式的线性组合 $d(x) = s(x)a(x) + t(x)b(x)$. (注意, 系数不一定是整数.)
   (c) 对 $x^{18} - 1$ 和 $x^{33} - 1$ 做 (a) 和 (b).
3. 在 $\mathbf{Q}$ 中求出 $2x^3 + 6x^2 - x - 3$ 和 $x^4 + 4x^3 + 3x^2 + x + 1$ 的最大公因子.
4. 假定多项式的系数是在 $\mathbf{Z}_3$ 中, 做习题 3.
5. 证明: $x^3 + x + 1$ 是模 5 不可约.
6. 在 $\mathbf{Z}_3$ 中对下列多项式进行因子分解:
   (a) $x^2 + x + 1$,　(b) $x^3 + x + 2$,　(c) $2x^3 + 2x^2 + x + 1$,
   (d) $x^4 + x^3 + x + 1$,　*(e) $x^4 + x^3 + x + 2$.
7. 在有理系数多项式环中列出 $x^4 - 1$ 的全部因子 (相伴除外). 证明: $x^4 - 1$ 的每个因子与你所列出的某个因子相伴.
8. 分别对 $x^6 - 1$ 和 $x^8 - 1$ 做习题 7.
9. 证明: $\mathbf{Z}$ 上两个多项式形式 $q(x)$ 和 $r(x)$ 表示 $\mathbf{Z}_p$ 上同一个函数当且仅当

$$(x^p - x) \mid [q(x) - r(x)].$$

（提示：利用 3.2 节习题 6.）

10. 证明：域上多项式的任意有限集合有一个最大公因子，它是已知集合中所有多项式的线性组合.
11. (a) 证明：域上任意两个已知多项式的所有公倍数组成的集合是一个理想.
    (b) 证明：多项式有最小公倍数；通过求 x^2+3x+2 和 $(x+1)^2$ 的最小公倍数加以说明.
12. 设给定 F 上的多项式 $p(x)$ 具有性质：$p(x)|a(x)b(x)$，总可以推出或者 $p(x)|a(x)$ 或者 $p(x)|b(x)$，证明 $p(x)$ 在 F 上是不可约的.
13. 证明：如果已知多项式 $p(x)$ 适合：任何其他多项式或者与 $p(x)$ 互素或者可被 $p(x)$ 整除，那么 $p(x)$ 是不可约的.
14. 设 $m(x)$ 是不可约多项式的幂，证明：由 $m(x)|a(x)b(x)$ 可推出或者 $m(x)|a(x)$，或者对某个 e，有 $m(x)|(b(x))^e$.
15. 设 $h(x)$ 与 $f(x)$ 和 $g(x)$ 两个多项式互素，证明：$h(x)$ 与 $f(x)g(x)$ 互素.
16. 证明：如果 $h(x)$ 与 $f(x)$ 互素，并且 $h(x)|f(x)g(x)$，则 $h(x)|g(x)$.
17. 证明：如果 $f(x)$ 和 $g(x)$ 是 $F[x]$ 中互素的多项式，并且 F 是 K 的子域，那么 $f(x)$ 和 $g(x)$ 在 $K[x]$ 中也是互素的.

*18. 证明：如果两个有理系数多项式具有公共实根，那么它们具有非常数公因子（也是有理系数多项式).

19. 下面给出有理系数多项式的某些集合. 这些集合中哪一些是理想？当集合是理想时，找出该集合中次数最低的多项式.
    (a) 满足 $b(3)=b(5)=0$ 的所有 $b(x)$；
    (b) 满足 $b(3)\neq 0$ 和 $b(2)=0$ 的所有 $b(x)$；
    (c) 满足 $b(3)=0, b(6)=b(7)$ 的所有 $b(x)$；
    (d) 使得 $b(x)$ 的某个幂可被 $(x+1)^4(x+2)$ 整除的所有 $b(x)$.
20. 设 S 是 F 上多项式的任意集合，它包含其中任意两个元素之差，并且当它包含任意 $b(x)$，就必包含 $xb(x)$ 和 $ab(x)$，这里 a 是 F 中任意常数，证明 S 是一个理想.

*3.9　其他唯一因子分解整环

考虑有理数域 $\mathbf{Q}$ 上关于两个未定元的多项式形式构成的整环 $\mathbf{Q}[x,y]$. $a(x,y)=x$ 和 $b(x,y)=y^2+x$ 的公因子只能是 1 及其相伴，但是不存在多项式 $s(x,y)$ 和 $t(x,y)$，满足关系式 $xs(x,y)+(y^2+x)t(x,y)=1$，这因为不管怎样选取 s 和 t，多项式 $xs+(y^2+x)t$ 总没有非零常数项. 类似地，在整系数多项式环 $\mathbf{Z}[x]$ 中，2 和 x 的最大公因子为 1，而关系式 $2s(x)+xt(x)=1$ 无解. 于是这两个整环中定理 12 都不成立.

然而我们可以证明，上述两种情况分解成素因子是可能的而且是唯一的（定理 14 成立).

定义 满足下列条件的整环称为唯一因子分解整环 (有时称为高斯整环):

(i) 非单位的任意元素可分解成素因子;

(ii) 除了相差次序和单位因子外, 这种因子分解是唯一的.

我们的主要结果是, 如果 G 是任意唯一因子分解整环, 那么 G 上任意多项式形式的整环 $G[x_1, \cdots, x_n]$ 同样是唯一因子分解整环. 对 n 用归纳法, 显然可把问题归结为关于单个未定元的 $G[x]$ 的情形, 我们将考虑这种情形.

首先, 我们把 G 嵌入 F 中, $F = Q(G)$ 为 G 的形式商构成的域(2.2 节定理 5), 并同 $G[x]$ 一起考虑 $F[x]$. 我们可以典型地把 G 想象为整数环, 相应地把 F 想象为有理数域.

其次, $F[x]$ 的多项式如果满足下列条件, 我们就称它为*本原多项式*: (i) 它的系数在 G 中 ("整数"), (ii) 它的所有系数没有除 G 中单位外的公因子. 例如 $3 - 5x^2$ 是本原多项式, $3 - 6x^2$ 就不是.

引理 1(高斯) 两个本原多项式的乘积是本原多项式.

证明 记

$$\sum_k c_k x^k = \sum_i a_i x^i \cdot \sum_j b_j x^j,$$

如果它不是本原多项式, 那么 G 中某素元素 p 将整除每个 c_k. 设 a_m 和 b_n 分别是 $\sum\limits_i a_i x^i$ 和 $\sum\limits_j b_j x^j$ 中第一个不能被 p 整除的系数 (它们确实存在, 因为这两个多项式都是本原的). 那么乘积的系数 c_{m+n} 的计算公式 (3) 给出

$$a_m b_n = c_{m+n} - [a_0 b_{m+n} + \cdots + a_{m-1} b_{n+1} + a_{m+1} b_{n-1} + \cdots + a_{m+n} b_0],$$

因为上式右边所有项都能被 p 整除, 所以乘积 $a_m b_n$ 能被 p 整除. 这就推出 p 必出现在 a_m 或者 b_n 的唯一因子分解式之中, 这与选取 a_m 和 b_n 为不能被 p 整除相矛盾.

引理 2 $F[x]$ 的任意非零多项式 $f(x)$ 可以写成 $f(x) = c_f f^*(x)$, 其中 c_f 在 F 中, $f^*(x)$ 是本原多项式. 此外, 对于给定的 $f(x)$, 常数 c_f 和本原多项式 $f^*(x)$ 除了相差一个可能的 G 的单位因子外是唯一的.

证明 首先记

$$f(x) = \frac{b_0}{a_0} + \frac{b_1}{a_1} x + \cdots + \frac{b_n}{a_n} x^n, \quad a_i, b_i \in G(\text{“整数”}),$$

设 $c = \dfrac{1}{a_0 a_1 \cdots a_n}$, 我们有 $f(x) = cg(x)$, 其中 $g(x)$ 的系数都在 G 中. 现在令 c' 是 $g(x)$ 的所有系数的最大公因子 (这是存在的, 因为 G 中唯一因子分解定理成立). 显然, $f^*(x) = \dfrac{g(x)}{c'}$ 是本原的, 并且 $f(x) = (cc')f^*(x)$, 取 $c_f = cc'$, 这就是引理中的第

一个结论.

为了证明 c_f 和 f^* 的唯一性, 只须证明 f^* 除了相差 G 的单位因子外是唯一的. 为此假定 $f^*(x)=cg^*(x)$, 其中 $f^*(x)$ 和 $g^*(x)$ 都是本原多项式, 并且 $c\in F$. 记 $c=\dfrac{u}{v}$, 其中 $u,v\in G$ 并且互素, 因此 $ug^*(x)=vf^*(x)$. 那么 v 就是 $ug^*(x)$ 的所有系数的公因子, 因为 u 和 v 互素, 所以 v 整除 $g^*(x)$ 的每个系数. 但是 $g^*(x)$ 是本原的, 因此 v 是 G 的单位. 由对称性, u 也是一个单位, 所以 $\dfrac{u}{v}$ 是 G 的单位. 这就完成了证明.

引理 2 的常数 c_f 称为 $f(x)$ 的容度, 除相差 G 中相伴元素外它是唯一的.

引理 3 *如果在 $G[x]$ 中或者甚至在 $F[x]$ 中有 $f(x)=g(x)h(x)$, 那么 $c_f\sim c_gc_h$, 并且 $f^*(x)\sim g^*(x)h^*(x)$, 这里"$\sim$"表示 $G[x]$ 中的相伴关系.*

证明 根据引理 1, $g^*(x)h^*(x)$ 是本原多项式, 显然它还是 $f^*(x)$ 的某常数倍. 根据引理 2, 两者仅相差 G 的一个单位因子 u(所以两者是相伴), 因此 $c_f=u^{-1}c_gc_h$.

证毕

这个引理的一个推论是, 如果 $f(x)$ 在 $G[x]$ 中, 并且它在 $F[x]$ 中是可约的, 那么 $f(x)=uc_fg^*(x)h^*(x)$. 这就给出下面关于定理 10 的推论的一个推广.

定理 15 *如果一个整系数多项式能分解成有理系数多项式之积, 那么它一定能分解成同次数的整系数多项式.*

更重要的是, 由引理 3, 在 $G[x]$ 中任意 $f(x)$ 的因子分解式分离成两个独立的部分: 一个是它的"容度" c_f 的分解, 一个是它的"本原部分" $f^*(x)$ 的分解. 前者相当于 G 的因子分解, 因此根据假设这种分解是可能的而且是唯一的. 根据引理 3, 后者本质上等价于 $F[x]$ 中的分解, 由定理 14, 这种分解是可能的而且是唯一的. 这就提出了

引理 4 *如果 G 是唯一因子分解整环, 那么 $G[x]$ 也是唯一因子分解整环.*

证明 由引理 2, 任何多项式 $f(x)$ 可分解成 $f(x)=c_ff^*(x)$, 因此 $G[x]$ 中的素元素 $f(x)$ 必然有因子 c_f 或者 f^* 中的一个是 $G[x]$ 的单位. 于是 $G[x]$ 的素元素分为两种类型: 一类是 G 的素元素 p, 一类是本原不可约多项式, 它不仅在 $G[x]$ 中而且在 $F[x]$ 中都是不可约的 (定理 15).

现在考虑 $G[x]$ 中任意多项式 $f(x)$. 它在 $F(x)$ 中有一个因子分解, 因而它与 $G[x]$ 的某些本原不可约多项式的乘积相伴, 记作 $f(x)\sim q_1(x)\cdots q_m(x)$. 于是 $f(x)=dq_1(x)\cdots q_m(x)$, 其中 G 的元素 d 可分解成 G 的素因子 p_i 的积, 总之, $f(x)$ 可分解成

$$f(x)=p_1\cdots p_rq_1(x)\cdots q_m(x),$$

其中每个 p_i 是 G 的素元素, 每个 $q_j(x)$ 是 $G[x]$ 的本原不可约多项式.

出现在这个因子分解式中的多项式 $q_j(x)$ 除相差 G 的单位外是唯一确定的, 它

是作为 $F[x]$ 中的 $f(x)$ 的唯一不可约因子的本原部分. 因为 $q_j(x)$ 都是本原的, 所以乘积 $p_1\cdots p_r$ 实质上是 $f(x)$ 的唯一的容度 c_f. 因此 $p_1,\cdots,p_r$(实质上) 是 c_f 在给定整环 G 中的全部因子 (唯一的). 这就证明了 $G[x]$ 是唯一因子分解整环.

由引理 4 并对 n 用归纳法我们可得出结论

定理 16 *如果 G 是任意唯一因子分解整环, 那么 G 上每个多项式整环 $G[x_1,\cdots,x_n]$ 也是唯一因子分解整环.*

14.10 节中我们将举出一个整环, 它不是唯一因子分解整环, 在这个整环上, 不论定理 12 还是定理 14 都不成立 (参见 3.6 节习题 11(g)).

习　题

1. 把下列各式表示成 $\mathbf{Z}[x]$ 的本原多项式与一个常数的乘积:

$$3x^2+6x+9, \qquad \frac{x^2}{2}+\frac{x}{3}+7.$$

2. 列出 $6x^2+3x-3$ 在 $\mathbf{Z}[x]$ 中的全部因子.

*3. 叙述一个求 $\mathbf{Z}[x]$ 中的多项式 $f(x)$ 的全部线性因子 $ax+b$ 的系统方法.

4. 整数 n 取什么值时, $2x^2+nx-7$ 在 $\mathbf{Q}[x]$ 中是可约的.

5. 找出下面多项式在 $\mathbf{Q}[x]$ 中的全体素因子:

$$x^3-1001x^2-1, \quad x^4+50x^2+2.$$

6. 证明: 在唯一因子分解整环中的两个元素 a 和 b 总有最大公因子 (a,b) 和最小公倍数 $[a,b]$.

7. 证明: 在任意唯一因子分解整环中, $ab\sim(a,b)[a,b]$.

8. 像 3.8 节习题 15 和习题 16 所指出的 "互素" 元素的性质, 在每个唯一因子分解整环中成立吗?

9. 按照正文的记号, 直接证明:

   (a) 在 $G[x]$ 中, $c_ff^*(x)|c_gg^*(x)$ 当且仅当在 G 中 $c_f|c_g$ 并且在 $F[x]$ 中 $f^*(x)|g^*(x)$.

   (b) 用 (a) 证明: 在 $G[x]$ 中整除乘积 $a(x)b(x)$ 的 "素元素" 必整除 $a(x)$ 或整除 $b(x)$.

10. 证明: 如果在 $F[x]$ 中 $f(x)$ 与 $g(x)$ 互素, 那么 $yf(x)+g(x)$ 在 $F[x,y]$ 中是不可约的.

11. 在 $\mathbf{Q}[x,y]$ 中, 把下列各式分解成不可约因子的乘积, 并证明分解出的因子确实是不可约的:

    (a) x^3-y^3, (b) x^4-y^2, (c) x^6-y^6, (d) $x^7+2x^3y+3x^2+9y$.

12. 找出 $\mathbf{Z}_2[x,y]$ 中所有次数 $\leqslant 2$ 的不可约多项式.

13. 证明: 在 $\mathbf{Q}[x,y]$ 中, 方程

$$1=s(x,y)(x-2)+t(x,y)(x+y-3)$$

不存在多项式解.

14. 证明：对于 $F[x,y]$ 中的多项式 $f(x,y)$，如果存在一个替换 $x\to t^r, y\to t^s$，它产生一个多项式 $f(t^r,t^s)$ 在 $F[t]$ 中是不可约的，并假定 $f(t^r,t^s)$ 的次数是所有整数 mr 和 ns 的最大值 (其中 m,n 是出现在 f 中的某一项 x^my^n 的指数)，那么 $f(x,y)$ 在 $F[x,y]$ 就是不可约的.

*15. (克罗内克尔 (Kronecker)) 证明：如果在 $\mathbf{Z}[x]$ 中 $f(x)|g(x)$，那么对 $\mathbf{Z}$ 中每个 c 有 $f(c)|g(c)$. 从这个事实 (和 3.2 节的插值公式 (5)) 出发，提出一个以有限步可以求得 $\mathbf{Z}[x]$ 中任意多项式 $f(x)$ 的给定次数的全部因子的系统方法.

16. 设 D 是所有这样有理数的集合，它可以写成分数 $\frac{a}{b}$，其分母 b 与 6 互素. 证明：D 是唯一因子分解整环.

*3.10 爱森斯坦不可约判别准则

显然，方程 $x^n=1$，当 n 为奇数时，除了 $x=1$ 外没有有理根. 由此得出 x^n-1 在 $\mathbf{Q}$ 上除了 $x-1$ 外没有首一线性因子. 但是这还不能证明商

$$\phi(x)=\frac{x^n-1}{x-1}=x^{n-1}+x^{n-2}+\cdots+x+1 \tag{13}$$

是不可约的，事实上，这个多项式，除 n 为素数外是可约的.

我们现在证明，如果 $n=p$ 是素数，那么由 (13) 定义的分圆多项式 $\phi(x)$ 是不可约的，因此 $x^p-1=(x-1)\phi(x)$ 给出 x^p-1 的 (唯一) 因子分解式 (分解成首一不可约因子). 这个结果将从下面关于不可约性的充分条件推出，这个定理是由爱森斯坦 (Eisenstein) 给出的.

定理 17 对于给定的素数 p，设

$$a(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_0$$

是一整系数多项式，如果

$$a_n\not\equiv 0 (\bmod p),\ a_{n-1}\equiv a_{n-2}\equiv\cdots\equiv a_0\equiv 0(\bmod p), a_0\not\equiv 0(\bmod p^2),$$

那么 $a(x)$ 在有理数域上是不可约的.

证明 假定可能有一个因子分解 $(n=m+k)$

$$a(x)=(b_mx^m+b_{m-1}x^{m-1}+\cdots+b_0)(c_kx^k+c_{k-1}x^{k-1}+\cdots+c_0),$$

根据定理 15，我们可以假定这两个因子都是整系数的，即 b_i, c_j 为整数. 因为 $a_0=b_0c_0$，所以假设中的第三个条件 $a_0\not\equiv 0(\bmod p^2)$ 意味着 b_0 和 c_0 不能同时被 p 整除. 固定一种情况，我们假设 $b_0\not\equiv 0(\bmod p)$ 而 $c_0\equiv 0(\bmod p)$. 但是 $b_mc_k=a_n\not\equiv$

$0 \pmod p$), 故 $c_k \not\equiv 0 \pmod p$. 选取最小的指标 $r(r \leqslant k)$ 使得 $c_r \not\equiv 0 \pmod p$, 而 $c_{r-1} \equiv \cdots \equiv c_0 \equiv 0 \pmod p$, 那么

$$a_r = b_0c_r + b_1c_{r-1} + \cdots + b_rc_0 \equiv b_0c_r \pmod p.$$

但是, 因为 p 是素数, 所以由 $b_0 \not\equiv 0$ 和 $c_r \not\equiv 0 \pmod p$ 得出 $a_r \not\equiv 0 \pmod p$. 根据假设, 这样的系数 a_r 只可能是 a_n, 故 $r = n$. 这表明第二个因子的次数必须是 n, 所以多项式 $f(x)$ 确实是不可约的. 证毕

当 $n = p$ 时, 这个判别准则可应用于多项式 (13), 这时 (13) 式给出分圆多项式

$$\phi(x) = \frac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \cdots + x + 1. \tag{13'}$$

实际上, 爱森斯坦判别准则还不能直接用于 (13′), 不过这可以做一个简单的变量替换 $y = x - 1$, 并由二项式展开得到

$$\frac{x^p - 1}{x - 1} = \frac{(y+1)^p - 1}{y} = y^{p-1} + py^{p-2} + \frac{p(p-1)}{1 \cdot 2}y^{p-3} + \cdots + p.$$

出现在上式右边的二项系数都是可被素数 p 整除的整数, 这是因为 p 作为因子出现在每个系数的分子中, 并且不能被分母中比它小的整数消去. 这样, 作为 y 的多项式满足爱森斯坦判别准则的假设条件, 因此是不可约的, 原来 (13′) 式的分圆多项式 $\phi(x)$ 仍保持这种不可约性.

习 题

1. 下列多项式中哪些在有理数域上是不可约的:

   $x^3 + 2x^2 + 4x + 2$, $x^3 + 2x^2 + 2x + 4$, $x^7 - 47$, $x^4 + 15$.

2. 用爱森斯坦判别准则证明: $x^2 + 1$ 在有理数域上是不可约的.
3. 证明: 如果 $f(x)$ 在域 F 上是不可约的, 那么对 F 中的任意 a, $f(x+a)$ 还是不可约的.
4. 证明: 如果 n 次 $(n > k)$ 多项式 $f(x)$ 满足条件 $a_n \not\equiv 0, a_k \not\equiv 0, a_{k-1} \equiv \cdots \equiv a_0 \equiv 0 \pmod p$, 且 $a_0 \not\equiv 0 \pmod{p^2}$, 那么 $f(x)$ 有一个次数至少为 k 的不可约因子.

*5. 证明: 如果 $2n+1$ 次奇次多项式 $f(x)$ 满足条件 $a_{2n+1} \not\equiv 0 \pmod p, a_{2n} \equiv \cdots \equiv a_{n+1} \equiv 0 \pmod p, a_n \equiv a_{n-1} \equiv \cdots \equiv a_0 \equiv 0 \pmod{p^2}, a_0 \not\equiv 0 \pmod{p^3}$, 那么 $f(x)$ 是不可约的.

6. (a) 设 $f(x)$ 为整系数首一多项式, 证明: $f(x)$ 对于模 p 的不可约性蕴含着它在 $\mathbf{Q}$ 上的不可约性.

   (b) 证明: $\mathbf{Z}$ 上的 $f(x)$ 的每个因子在模 p 之下必可化为 $\mathbf{Z}_p$ 上相同次数的因子.

   (c) 用这个办法 (用小的素数 p) 检验下列多项式在 $\mathbf{Q}$ 上的不可约性:

   $x^3 + 6x^2 + 5x + 25$, $x^3 + 6x^2 + 11x + 8$, $x^4 + 8x^3 + x^2 + 2x + 5$.

7. (a) 设 $F[t]$ 是关于未定元 t 的全体多项式的整环, 对于系数在 $F[t]$ 上的多项式 $f(x)$, 叙述并证明类似于爱森斯坦判别准则的定理. (提示: 用 t 代替 p.)

   (b) 用此定理证明 $x^3 + 3t^2x^2 + 2tx^2 + t^4x + 7t + t^2$ 在 $F[t, x]$ 中是不可约的.

*3.11 部 分 分 式

多项式的唯一因子分解定理可以用于有理函数上,以便得到某些简化的表达式,例如在积分学中用到的部分分式的展开. 现在我们就对此进行讨论, 这一节涉及到的多项式和有理分式都假定它们的系数是在某一固定的域 F 上.

首先考虑这样的有理式$\dfrac{b(x)}{a(x)}$, 它的分母分解成互素的因子 $c(x)$ 和 $d(x)$, 即 $a(x) = c(x)d(x)$. 由定理 12 给出多项式 $s(x)$ 和 $t(x)$ 适合 $1 = sc + td$, 因此

$$\frac{b(x)}{c(x)d(x)} = \frac{s(x)b(x)}{d(x)} + \frac{t(x)b(x)}{c(x)}. \tag{14}$$

这一结果可叙述成

引理 1 *一个有理分式, 如果它的分母是两个互素多项式 $c(x)$ 和 $d(x)$ 的乘积, 那么它可以表示成分母分别为 $c(x)$ 和 $d(x)$ 的两个商式之和.*

如果分母 $a(x)$ 是一个幂 $a(x) = [c(x)]^m, m > 1$, 那么这个方法还不能直接应用. 改换另法, 按照辗转相除法, 用 $c(x)$ 去除分子, 有

$$b(x) = q_0(x)c(x) + r_0(x),$$

然后再用 $c(x)$ 去除商 $q_0(x)$ 得

$$q_0(x) = q_1(x)c(x) + r_1(x),$$

两个等式合起来便得

$$b(x) = q_1(x)[c(x)]^2 + r_1(x)c(x) + r_0(x).$$

重复这一过程 (注意, 这个说法隐含了归纳过程), 采用缩写记号我们得到①

$$b(x) = q_{m-1}c^m + r_{m-1}c^{m-1} + \cdots + r_1c + r_0, \tag{15}$$

这里每个多项式 $r_i = r_i(x)$, 如果不为零, 则它的次数低于 $c(x)$ 的次数. 现在有理分式 $\dfrac{b(x)}{a(x)}$ 变成

$$\frac{b}{c^m} = q_{m-1} + \frac{r_{m-1}}{c} + \frac{r_{m-2}}{c^2} + \cdots + \frac{r_1}{c^{m-1}} + \frac{r_0}{c^m}. \tag{16}$$

这就证明了

① 这同 1.5 节习题 11 关于整数的十进制小数展开式相类似.

引理 2 *以幂 $[c(x)]^m$ 为分母的有理式可以表示成一个多项式加上一些有理分式之和, 每个有理分式的分母是 $c(x)$ 的幂, 分子的次数低于 $c(x)$ 的次数.*

综合这些结果, 可把任意给定的分母 $a(x)$ 分解成首一不可约多项式的乘积. 如果把相同的不可约因子归在一起, 我们有

$$a(x) = a_0[p_1(x)]^{m_1}[p_2(x)]^{m_2}\cdots[p_k(x)]^{m_k}, \tag{17}$$

其中指数 m_i 为整数. 任意两个不同的首一不可约多项式 $p_1(x)$ 和 $p_2(x)$ 当然互素, 因此幂 $[p_1(x)]^{m_1}$ 和 $[p_2(x)]^{m_2}$ 除了单位外没有公因子, 所以是互素的. 因此可以应用引理 1 把分母分解成其中一个因子是 $c_1(x) = [p_1(x)]^{m_1}$, 而另一个因子是 (17) 式除去 $[p_1(x)]^{m_1}$ 后余下的部分. 重复进行下去, 就可把 $\dfrac{b}{a}$ 表为一些分式的和, 每个分式的分母是 $[p_i(x)]^{m_i}$. 最后还可用 (16) 式把每个分式进一步化简.

定理 18 *任意有理分式 $\dfrac{b(x)}{a(x)}$ 可以表示成一个 x 的多项式加上形为 $\dfrac{r(x)}{[p(x)]^m}$ 的 ("部分") 分式之和, 这里 $p(x)$ 为不可约多项式, $r(x)$ 的次数低于 $p(x)$ 的次数. 所出现的分母 $[p(x)]^m$ 是原来分母 $a(x)$ 的一切因子.*

如果要找到给定的有理函数 $\dfrac{b(x)}{a(x)}$ 的明显的部分分式表达式, 那么照定理 18 证明的每一步去做就可得到. 这样的证明称为"构造性"证明, 它总是可以用于有关对象的实际计算.

例如, 在有理数域 $\mathbf{Q}$ 上考虑 $\dfrac{x+1}{x^3-1}$. 分母是 $(x-1)(x^2+x+1)$, 第二个因子是不可约的, 由辗转相除法得到 $x^2+x+1=(x+2)\cdot(x-1)+3$. 用原来分式的分子 $(x+1)$ 乘这个方程, 我们得到

$$3(x+1) = (x+1)(x^2+x+1)-(x^2+3x+2)(x-1);$$

$$\frac{3(x+1)}{x^3-1} = \frac{x+1}{x-1} - \frac{x^2+3x+2}{x^2+x+1}.$$

所得的每个分式可以通过辗转相除法进一步化简[①]得出

$$\frac{3(x+1)}{x^3-1} = \frac{2}{x-1} - \frac{2x+1}{x^2+x+1}.$$

在实数域 $\mathbf{R}$ 上, 不可约多项式只能是线性多项式和满足 $b^2-4ac<0$ 的二次多项式 ax^2+bx+c. (这个命题将在 5.4 节定理 7 中证明.) 因此, 在 $\mathbf{R}$ 上任意有理函

① 把这个直接方法同微积分教材中常用的方法加以比较, 在那里, 我们必须解出出现在项 $\dfrac{A}{x-1}$ 和 $\dfrac{Bx+C}{x^2+x+1}$ 中的未知系数 A, B, C.

数可以表示成分母是线性多项式的幂和二次多项式的幂的分式之和. 这个事实在微积分学中用来证明：任何有理函数的不定积分可以通过“初等函数”(即代数函数、三角函数、指数函数以及它们的反函数) 来表示. 根据定理 18, 有理分式求积分时, 本质上可化为 $\dfrac{c}{(x+a)^m}$ 和 $\dfrac{c(x+d)}{(x^2+ax+b)^m}$ 两种类型的项之和求积分. 因此, 如果这两种类型的函数的积分可以通过初等函数表示 (这是可以做到的), 那么关于积分的命题就将被证明.

习　　题

1. 分解下式成部分分式 (在有理数域上)：

   (a) $\dfrac{3x+4}{x^2+3x+2}$,　(b) $\dfrac{1}{x^2-a^2}$,　(c) $\dfrac{1}{x^3+x}$,

   (d) $\dfrac{a^2}{x^3-a^3}$,　(e) $\dfrac{3}{x^4+5x^2+4}$,　(f) $\dfrac{3x-7}{(x-2)^2}$.

2. 分别在下列域上把 $\dfrac{4x+2}{x^3+2x^2+4x+8}$ 分解成部分分式：

   (a) 模 5 整数域 $\mathbf{Z}_5$,　(b) 有理数域 $\mathbf{Q}$.

3. 设 $a_0, a_1, \cdots, a_n$ 为不同元素, 证明：

$$\frac{1}{\prod_i (x-a_i)} = \sum_i \frac{C_i^{-1}}{x-a_i}, \quad \text{其中 } C_i = \prod_{j\neq i}(a_i-a_j).$$

   (提示：用拉格朗日插值公式展开 $p(a_i)=1$.)

4. 对 m 用归纳法, 证明 (13) 式.

5. 用归纳法给出定理 18 的详细证明.

6. (a) 证明：任意非多项式的有理式可以表示成一个多项式与另一个有理式之和, 这个有理式的分子是次数低于分母次数的多项式.

   (b) 这种表示是唯一的吗?

7. 如果限定所有分式 (包括部分分式) 的分子的次数低于相应的分母的次数, 证明下列引理和定理中的表达式是唯一的：

   (a) 在引理 1 中,　(b) 在引理 2 中,　(c) 在定理 18 中.

8. (a) 设 $(x-a)$ 不是 $f(x)$ 的因子, 证明

$$\frac{1}{(x-a)^r f(x)} = \frac{C}{(x-a)^r} + \frac{g(x)}{(x-a)^{r-1} f(x)},$$

其中 $C=\dfrac{1}{f(a)}$, 并且 $g(x)$ 是一个适当的多项式.

*(b) 对于分母可分解成线性因子的有理函数, 利用习题 8(a) 或习题 3 化成标准形式.

9. (a) 设 $p(x)$ 是不可约的, 证明：一个分式 $\dfrac{b(x)}{p(x)}$(b 与 p 互素) 的任何表示成分式之和的表达式, 必至少包含一个其分母可被 $p(x)$ 整除的分式. (这就意味着 $\dfrac{b(x)}{p(x)}$ 的进一步部分分式的分解是不可能做到的.)

(b) 对于 $\dfrac{b(x)}{[p(x)]^m}$ 能有相同的说法吗?

*10. 求下式之和：

$$\frac{1}{(x+1)(x+2)}+\frac{2}{(x+2)(x+4)}+\cdots+\frac{2^n}{(x+2^n)(x+2^{n+1})}.$$

*11. 建立表示任何有理数为“部分分式”之和的方法, 其中部分分式具有特殊形式 $\dfrac{a}{p^n}$(p 为素数, $0\leqslant a<p$). 例如 $\dfrac{1}{6}=\dfrac{1}{2}-\dfrac{1}{3}$.

*12. 假定 5.3 节的定理 6 成立, 证明：任意一个复有理函数的不定积分是由一个有理函数与复对数 $\ln(z+a_i)=\displaystyle\int\frac{\mathrm{d}z}{z+a_i}$ 的线性组合组成的和.

*13. 证明：在任意有序整环 D 上, 如果我们选取那些具有正的首项系数的多项式 (也就是在 (1) 式中 $a_n>0$) 作为“正”的多项式, 那么多项式环 $D[x]$ 就成为有序整环.

第4章　实　　数

4.1　毕达哥拉斯二难推论

抽象代数虽然相当多地强调了一般的域和整环所具有的大量性质, 但是实数域和复数域对于定量地描述我们生活的世界还是不可缺少的. 例如, 在代数和几何的关系中, 不仅在初等解析几何而且在进一步讨论向量和向量分析 (第 7 章) 中, 这两个域都是很重要的. 此外, 它们还具有独特的代数性质, 这些性质将在本书后几章中展示出来. 特别重要的是实数域**R**的*序*的完备性和复数域**C**的*代数*完备性. 我们在第 4 章至第 5 章叙述这些完备性及其代数含意.

希腊人研究实数用的是纯几何方法. 对他们来说, 一个数只不过是两个线段 a 和 b 的长度之比 $(a{:}b)$. 他们直接给出关于比的相等以及比的加法、乘法、减法和除法的几何构造. 全体实数构成*有序域*(2.4 节) 的公设在希腊人看来, 是由平面几何一系列公设 (包括平行公设) 证明的一组几何定理.

古希腊哲学家毕达哥拉斯 (Pythagoras) 知道, 正方形对角线长 d 与它的边长 s 之比 $r=\dfrac{d}{s}$ 一定满足方程

$$d^2=(rs)^2=r^2s^2=s^2+s^2 \quad (\text{毕达哥拉斯定理}). \tag{1}$$

因此他推理, *存在*一个 "数" r, 满足 $r^2=1+1=2$.

另一方面, 他发现 r 不能表示成两个整数的商 $r=\dfrac{a}{b}$, 这是因为 $\left(\dfrac{a}{b}\right)^2=2$ 意味着 $a^2=2b^2$, 根据素因子分解定理, 2 每次整除 a, 就恰有两次整除 a^2, 因此 2 整除 a^2 偶数次, 类似地, 2 整除 $2b^2$ 奇数次, 所以 $a^2=2b^2$ 没有整数解.

只要把不能表为整数之商的数设为*无理数*, 我们就能避开上述 "毕达哥拉斯二难推论".

类似的论证指出, 立方体 C 的对角线长与它的边长之比为 $\sqrt{3}$, C 的边长与具有一半体积的立方体的边长之比为 $\sqrt[3]{2}$, 这些都是无理数. 这些结果是 3.7 节定理 10 的特殊情况.

此外, π, e 及许多别的数都是无理数 (因此 π 不会恰好是 $\dfrac{22}{7}$, 甚至 3.141 6). 我们在第 14 章将证明, 绝大多数的实数不仅是无理数, 而且甚至不能满足任何一个代数方程 (与 $\sqrt{2}$ 不同). 为了回答 "什么是实数?" 这个基本问题, 我们要用到一些新的概念.

一个概念是连续性—— 如果实轴分成两段, 那么这两段必在公共边界点上相接. 第二个概念是, 有序的有理数域Q在实数域里稠密, 因此, 每个实数是一个或多个有理数序列 (例如精确到 n 位的有限小数逼近序列) 的极限. 这个概念还可表述为

$$如果\ x<y,\ 那么存在\ \frac{m}{n}\in\mathbf{Q},\quad 使得\ x<\frac{m}{n}<y. \tag{2}$$

实数的这个性质首先被希腊数学家欧多克斯 (Eudoxus) 发现. 欧多克斯把 $x=a:b$ 和 $y=c:d$ 都看作线段长度之比, 线段长度 a 的整数倍 na 可用几何方法做出, 他规定 $(a:b)=(c:d)$ 当且仅当对一切正整数 m 和 n,

$$由\ na>mb\ 推出\ nc>md,\quad 由\ na<mb\ 推出\ nc<md. \tag{3}$$

上述这两个概念可以合并成一个完备性公设, 由这个公设我们可以通过有序域Q的自然扩张来构造实数域. 这个"完备性"公设类似于整数的良序公设 (1.4 节), 二者都涉及无限集合的性质, 这种性质是非代数的. 我们将要看到, 这个完备性公设, 对于建立实数域的某些重要代数性质 (例如每个正数都有平方根) 是必需的.

习　题

1. 给出 $\sqrt{3}$ 是无理数的直接证明.
2. 证明: $\sqrt[n]{a}$ 是无理数, 除非整数 a 是某一整数的 n 次幂.
3. 证明: $\log_{10}3$ 是无理数. (提示: 利用对数定义.)
4. 证明: 如果 $a\neq 0$ 和 b 都为有理数, 那么, $au+b$ 为有理数当且仅当 u 为有理数.
5. 证明: $\sqrt{2}+\sqrt{5}$ 是无理数. (提示: 从 $x-\sqrt{2}=\sqrt{5}$ 两边平方出发, 找出一个以 $\sqrt{2}+\sqrt{5}$ 为根的多项式方程.)

*6. 证明: 定义为收敛级数 $\sum_{k=0}^{\infty}\frac{1}{k!}$ 的数 e 是无理数. (提示: 如果是有理数, 那么对某个 n, $(n!)$e 可以是整数.)

4.2 上界与下界

实数域可以最简单地被描述为具有下列性质的有序域, 即域中任意有界集合都有最大下界和最小上界. 我们现在就定义这两个概念, 它们类似于可除性理论中的最大公因子和最小公倍数的概念.

定义 设 S 为有序整环 D 中某些元素构成的集合, 如果 D 中元素 b(它本身不一定在 S 中) 使得对 S 中每个元素 x, 有 $b\geqslant x$, 则称 b 为 S 的上界. 对于 S 的上界 b, 如果 D 中没有比 b 小的元素是 S 的上界, 也就是说, 如果对任意 $b'<b$, S 中都存在一个 x 适合 $b'<x$, 那么 b 就是 S 的最小上界.

把上面定义中的 “>” 换成 “<”, “<” 换成 “>”, 可定义 S 的下界和最大下界的概念.

由定义直接可知, D 的子集 S 至多有一个最小上界, 并且至多有一个最大下界 (为什么?).

直观上, 把实数当作连续直线 (x 轴) 上的点来考虑, 并想象在这条直线上, 全体有理数密集地撒布在它们各自本来的位置上. 由此我们容易得出结论：每个实数 a 可定义为所有适合 $r < a$ 的有理数 $r = \dfrac{m}{n}(n > 0)$ 的集合 S 的最小上界. 例如, $\sqrt{2}$ 是大于所有适合 $m^2 < 2n^2$ 的比 $\dfrac{m}{n}(m > 0, n > 0)$ 的最小实数. 也就是数, 数 $\sqrt{2}$ 是适合 $m^2 < 2n^2$ 的正有理数 $\dfrac{m}{n}$ 的集合的最小上界.

用无限小数表示实数的普通表达式直接包含着把实数看作有理数集合的最小上界的概念. 例如我们可以把 $\sqrt{2}$ 写成最小小界(l.u.b.)和最大下界 (g.l.b.) 两种形式

$$\begin{aligned}\sqrt{2} &= \text{l.u.b.}(1.4,\ 1.41,\ 1.414,\ 1.4142, \cdots)\\ &= \text{g.l.b.}(1.5,\ 1.42,\ 1.415,\ 1.4143, \cdots).\end{aligned} \tag{4}$$

由熟悉的小数表达式的性质很容易看出, 每个正实数非空集合 T 有最大下界, 如下所述.

考虑把 T 的元素只取前 n 位的 n 位小数, 它们中间必有一个最小的元素, 这是因为只有有限个非负 n 位小数, 比 T 的任何给定的元素都小. 设这个最小的 n 位小数是 $k + 0.d_1d_2\cdots d_n$, 其中 k 为某整数, d_i 为数字. 最小的 $n+1$ 位小数其前 n 位与 $d_1d_2\cdots d_n$ 相同, 因此有形式 $k + 0.d_1d_2\cdots d_nd_{n+1}$, 这里添上一个数字 d_{n+1}. 所以上述构造定义了某一个无限小数

$$c = k + 0.d_1d_2d_3\cdots.$$

根据构造, c 就是 T 的下界 (因为 T 中没有一个 x 能比 c 的小数表达式小), 而且是最大下界 (任何比 c 大的小数就不再是 T 的下界了).

但是, 如果把实数定义成无限小数, 那么很难证明中学代数中所承认的事实：无限小数系统是有序域 ①.

习　　题

1. 证明：$x = 0.124\ 374\ 374\ 37\cdots$ 表示有理数. (提示：计算 $1000\cdot x - x$.)
2. 证明：$y = 1.236\ 723\ 67\cdots$ 表示有理数.

*3. 证明：像习题 1 和习题 2 那样的任意 “循环小数” 表示有理数. 对 “循环小数” 这个术语给出定义.

① 详细请看 J. F. Ritt, *Theory of Functions* (New York, Kings Grown Press, 1947). 困难在于, 两个不同的小数实际上是相等的, 例如 0.199 99$\cdots$=0.200 00$\cdots$.

*4. 反过来证明：任意有理数的小数表达式是“循环的”. (提示：把有理数表示成分数. 并证明, 在十进制下, 如果在分母去除分子过程中, 第 m 次同第 $m-k$ 次的余数相同, 那么所得的商数的数字中就分成 k 个数字一节无限地重复下去.)

*5. 在十二进制下, 习题 4 的结论成立吗？

6. 在以 3 的幂为分母的所有有理数组成的整环中, 求出 $\sqrt{2}$ 的三个逐次近似值.

7. 构造出两个不同的有理数集合, 它们都以 2 为最小上界.

*8. 序列 $2, \frac{3}{2}, \frac{17}{12}, \frac{577}{408}, \cdots$ 由递推公式

$$x_1 = 2, \quad x_{k+1} = \frac{x_k}{2} + \frac{1}{x_k}$$

定义.

(a) 证明：对 $k > 1$ 有 $x_k = \frac{m_k}{n_k}$, 其中 $m_k^2 = 2n_k^2 + 1$.

(b) 定义 $\varepsilon_k = x_k - \sqrt{2}$, 证明：$0 < \varepsilon_{k+1} < \frac{\varepsilon_k^2}{2\sqrt{2}}$.

(c) 证明：g.l.b.$(2, \frac{3}{2}, \frac{17}{12}, \cdots)=\sqrt{2}$.

4.3 实数公设

我们现在将用一组简短的公设来描述实数. 后面我们会看到 (定理 6), 这些公设唯一地 (精确到同构) 确定全体实数.

定义 *有序整环 D 是完备的当且仅当 D 的正元素的每个非空集合在 D 中有最大下界.*

实数公设　实数构成完备的有序域 $\mathbf{R}$.

我们根据这个公设得出的实数性质, 确实可以导出全体实数的一切熟知的性质, 包括像罗尔 (Rolle) 定理那样的结果, 这个定理在微积分学中, 对于泰勒 (Taylor) 定理的证明或其他方面是基本的.

但是我们只限于讨论几个简单的应用.

定理 1 *在实数域 $\mathbf{R}$ 中, 每个具有下界的非空子集 S 有最大下界, 每个具有上界的非空子集 T 有最小上界.*

证明 假设 S 有下界 b. 如果把 $1-b$ 加在 S 的每个数 x 上, 则得到正数 $x-b+1$ 的集合 S'. 根据实数公设, 这个集合 S' 具有最大下界 c'. 因此数 $c=c'+b-1$ 是原来集合 S 的最大下界, 这很容易验证.

对偶地, 如果集合 T 有上界 a, 则 T 的所有元素的负元素 $-y$ 组成的集合具有下界 $-a$, 因此根据上述证明, 它有最大下界 b^*. 那么可以证明, 数 $a^* = -b^*$ 就是已知集合 T 的最小上界. 证毕

实数公设保证全体实数构成有序域 $\mathbf{R}$, 所以 2.6 节定理 18 的推论 2 指出, $\mathbf{R}$ 必包含同构于有理数域 $\mathbf{Q}$ 的子域. 因为在第 2 章里, $\mathbf{Q}$ 仅在同构意义下定义, 所以我

们也可同样假定实数域 $\mathbf{R}$ 包含所有有理数, 因而包含所有整数. 这个约定使上述公设适应于习惯用法, 并可使我们证明下面的实数性质 (常称为阿基米德定律).

定理 2 *在所有实数组成的域 $\mathbf{R}$(由实数公设定义的) 中, 对任意两个数 $a>0$ 和 $b>0$, 存在整数 n 使得 $na>b$.*

证明 假定对于两个特定的实数 a 和 b, 上述结论是错误的, 因而对每个 n, 有 $b\geqslant na$. 则所有倍数 na 的集合 S 有上界 b, 从而它也有最小上界 b^*. 所以对每个 n, $b^*\geqslant na$, 因此对每个 m 也有 $b^*\geqslant(m+1)a$. 这推出 $b^*-a\geqslant ma$, 所以 b^*-a 是 a 的所有倍数的集合 S 的上界, 但是它小于已知的最小上界, 得出矛盾.

推论 *对已知实数 a 和 $b(b>0)$, 总存在整数 q 适合 $a=bq+r, 0\leqslant r<b$.*

这是除法算式的推广, 证明留给读者.

这样建立的"阿基米德性质"可以证明欧多克斯条件 (参见 4.1 节 (3)) 是合理的.

定理 3 *任意两个实数 c 和 d 之间 $(c>d)$, 存在有理数 $\frac{m}{n}$ 适合 $c>\frac{m}{n}>d$.*

像前面那样, 这个定理由"实数构成完备的有序域"的公设就可证明. 根据假设, $c-d>0$, 所以由阿基米德定律有正整数 n 使得 $n(c-d)>1$, 即 $\frac{1}{n}<c-d$. 现在设 m 为适合 $m>nd$ 的最小整数, 那么 $\frac{m-1}{n}\leqslant d$, 因此

$$\frac{m}{n}=\frac{m-1}{n}+\frac{1}{n}<d+(c-d)=c.$$

因为 $\frac{m}{n}>d$, 这就完成了证明.

我们可以对上面的证明直观说明如下. 具有固定分母 n 的不同分数0, $\pm\frac{1}{n}$, $\pm\frac{2}{n},\cdots$, 以长度 $\frac{1}{n}$ 为间隔沿实轴隔开. 为确保一个这样的点能落在 c 和 d 之间, 我们只需使间隔 $\frac{1}{n}$ 小于已知的差 $c-d$.

这个定理可用来正式地证明像(4) 式那样把实数表示成有理数集合的最小上界的直观想法.

推论 *每个实数都是某有理数集合的最小上界.*

证明 对于已知实数 c, 设 S 表示所有有理数 $\frac{m}{n}\leqslant c$ 的集合, 那么 c 是 S 的上界, 根据定理 3, 没有比 c 小的实数 d 可以是 S 的上界, 因此 c 是 S 的最小上界.

习 题

1. 证明: 不存在这样的有序整环 D, 其中每个非空集合具有最小上界. (提示: D 本身可以没有上界.)

*2. 证明: 有序整环 $\mathbf{Z}$ 是完备的.

3. 用几何语言叙述关于实轴的点的公设, 它断言: 有界集合具有最小上界和最大下界 (用词 "左" 和 "右").
4. 分别指出下列有理数集合的最小上界:
(a) $\frac{1}{3}, \frac{4}{9}, \frac{13}{27}, \frac{40}{81}, \cdots$, (b) $\frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \frac{15}{16}, \cdots$.
5. 设集合 S 具有最小上界 a^* 和最大下界 b^*.
(a) 详细证明: 为什么所有数$-3x$(x在S中)的集合具有最小上界$-3b^*$ 和最大下界$-3a^*$.
(b) 用同样的方法找出, 所有数 $x+5$(x在S中) 的集合的最小上界和最大下界.
6. 在习题 5 的条件下, 并假设 $b^* > 0$, 分别找出下列集合的最小上界.
(a) 所有数 $7x+2$(x在S中) 的集合,
(b) 所有数 $\frac{1}{x}$($x \neq 0$, 在S中) 的集合.
7. 设 S_1 和 S_2 为分别具有最小上界 b_1 和 b_2 的实数集合, 找出下列集合的最小上界:
(a) 所有和 $s_1 + s_2$($s_1 \in S_1, s_2 \in S_2$) 的集合 $S_1 + S_2$,
(b) 或属于 S_1 或属于 S_2 的所有元素的集合.
8. 集中列出一组完整的实数公设.
*9. 构造一组正实数公设. (提示: 参见 2.5 节)
10. 证明: 在有序域中, 一个元素 a^* 是集合 S 的最小上界当且仅当 (i) 对一切 $x \in S$, $x \leqslant a^*$; (ii) 对域中每个正的 e, S 中都有一个 x 适合

$$|x - a^*| < e.$$

11. 证明: 任意两个实数 c 和 d 之间 ($c < d$), 存在有理数的立方 $\left(\frac{m}{n}\right)^3$ 适合 $c < \left(\frac{m}{n}\right)^3 < d$. 对于有理数的平方, 这个结论还正确吗?
12. 设 $n > 1$ 是整数, 证明: 任意两个实数 c 和 d 之间 ($c > d$), 存在形为 $\frac{m}{n^k}$ 的有理数, 其中 m 和 k 为适当的整数.
13. 设 a, b, c 和 d 为完备的有序域的正元素, 证明: $\frac{a}{b} = \frac{c}{d}$ 当且仅当欧多克斯的条件 (3) 式成立.
14. 详细证明定理 2 的推论.

4.4 多项式方程的根

我们现在将指出怎样利用最小上界的存在性来证明实数系 $\mathbf{R}$ 的各种性质, 首先包括像 $x^2 = 2$ 这样方程解的存在性.

定理 4 *如果 $p(x)$ 为实系数多项式, $a < b$, 并且 $p(a) < p(b)$, 那么对满足 $p(a) < C < p(b)$ 的每个常数 C, 方程 $p(x) = C$ 在 a 和 b 之间有根.*

几何上, 定理的假设意味着 $y = p(x)$ 的曲线与水平直线 $y = p(a)$ 在 $x = a$ 处相交, 并与水平直线 $y = p(b)$ 在 $x = b$ 处相交; 定理的结论是说: 曲线也必与每条

中间的水平直线 $y = C$ 在某点相交[①], 这点的 x 坐标在 a 和 b 之间.

证明依赖于下面两个引理.

引理 1 对任意实数 x 和 h, 我们有

$$p(x+h) - p(x) = hg(x,h),$$

其中 $g(x,h)$ 是只依赖于 $p(x)$ 的多项式.

证明 (参见 3.2 节定理 3) 根据二项定理, 对于 $p(x)$ 的每个单项 $a_k x^k$, 这是正确的. 现在对 k 求和并且提出公因子 h, 我们便得到所需要的结论.

引理 2 对于已知的 a,b 和 $p(x)$, 存在实常数 M, 使得满足 $a \leqslant x \leqslant b$, $a \leqslant x+h \leqslant b$ 的一切 x 和一切正的 h, 有

$$|p(x+h) - p(x)| \leqslant Mh.$$

证明 根据引理 1, 只须证明当 $x \leqslant |a| + |b|, |h| \leqslant |b-a|$ 时, 有 $|g(x,h)| \leqslant M$. 但是, 如果我们把 $g(x,h)$ 的每项用它的绝对值代替, 根据 1.3 节的公式 (3), 则使 $|g(x,h)|$ 增大或保持不变. 如果我们再分别用 $|a| + |b|$ 和 $|b-a|$ 代替 $|x|$ 和 $|h|$, 那么又使得到的结果增大或保持不变. 然而, 这个替换给我们一个只依赖于 $p(x)$ 的系数和区间 $a \leqslant x \leqslant b$ 的实常数 M.

有了引理 2, 我们准备证明定理 4. 设 S 表示 a 和 b 之间满足 $p(x) \leqslant C$ 的实数的集合. 因为 $p(a) < C$, 所以 S 是不空的, 并且它以 b 为上界. 因此 S 有实的最小上界 c, 我们来证明 $p(c) = C$.

为此目的, 显然只须排除 $p(c) < C$ 和 $p(c) > C$ 两种可能. 但是, 根据引理 2 由 $p(c) < C$ 可推出, 对 $h = \dfrac{C - p(c)}{M}$, $p(c+h) \leqslant C$, 因此 $(c+h) \in S$. 这与 c 是 S 的上界的定义矛盾. (引理 2 可以应用, 是因为 $c+h \geqslant b$ 显然是不可能的).

现在剩下 $p(c) > C$ 这种可能. 但是在这种情形下, 再根据引理 2, 对一切正的 $h \leqslant \dfrac{p(c) - C}{2M}$, 有 $p(c-h) > C$, 这与 c 是 S 的最小上界的定义矛盾: $c - \dfrac{p(c) - C}{2M}$ 将给出比 c 小的上界. 因此只留下 $p(c) = C$. 证毕

根据这个定理我们容易证明:

推论 1 如果 $p(x)$ 为正系数多项式, 而且没有常数项, $C > 0$, 那么 $p(x) = C$ 有正实根.

推论 2 如果 $p(x)$ 为奇次多项式, 那么对每个实数 $C, P(x) = C$ 有实根.

定理 4 没有给出 $p(x) = C$ 的小数形式的根的实际计算方法, 但这是容易做到的. 例如, 我们可以设 $c_1 = \dfrac{a+b}{2}$, 则或 $p(c_1) = C$, 或 $p(c_1) > C$, 或 $p(c_1) < C$. 在

① 数学分析中有一个一般性的定理, 它断言这个结论不仅对于多项式函数 $p(x)$ 成立, 而且对于任意连续函数也成立.

第一种情况下, 方程的根就找到了; 在第二、三种情形下, 分别在区间 $a \leqslant x \leqslant c_1$ 和 $c_1 \leqslant x \leqslant b$ 中有根, 这个区间长度为原来区间的一半. 重复这一过程便可得到 $p(x)=C$ 的根的任何精度的近似值.

如果采用线性内插法, 且令

$$c_1 = a + \frac{[C-p(a)](b-a)}{p(b)-p(a)},$$

则收敛得更快.

其他计算方程根的有效方法在数学分析里讨论. 例如, 若 $|x|<1$, 我们可以运用无穷级数

$$\sqrt{1+x} = 1 + \frac{1}{2}x + \frac{1}{2}\Big(-\frac{1}{2}\Big)\frac{x^2}{2!} + \frac{1}{2}\Big(-\frac{1}{2}\Big)(-\frac{3}{2})\frac{x^3}{3!} + \cdots. \tag{5}$$

附录 三次方程的三角解法

在三次方程

$$a_3x^3 + a_2x^2 + a_1x + a_0 = 0, \quad a_3 \neq 0 \tag{6}$$

的情况下, 实根可按下法求得. 我们用 a_3 去除方程各项把 (6) 化简成 $a_3=1$ 的情形. 现在作变量替换 $x = y - \frac{a_2}{3}$, 并移动常数项, 把 (6) 化为

$$y^3 + py = q. \tag{7}$$

当 $p=0$ 时, 答案立即可得.

否则, 令 $y=hz$, 并用 k 乘 (7) 的各项, 其中 $h=\sqrt{\frac{4|p|}{3}}, k=\frac{3}{h|p|}$, 我们可把 (7) 化为下列两个方程之一:

$$4z^3+3z=C \quad \text{或} \quad 4z^3-3z=C. \tag{8}$$

为了解第一个方程, 我们可用熟悉的三角恒等式

$$\mathrm{sh}3\theta = 4\mathrm{sh}^3\theta + 3\mathrm{sh}\theta,$$

因此

$$z = \mathrm{sh}\Big(\frac{1}{3}\mathrm{arsh}\ C\Big). \tag{9a}$$

为了解第二个方程, 当 $C \geqslant 1$, 我们利用类似的公式

$$\mathrm{ch}3\theta = 4\mathrm{ch}^3\theta - 3\mathrm{ch}\theta,$$

得到

$$z = \mathrm{ch}\Big(\frac{1}{3}\mathrm{arch}\ C\Big). \tag{9b}$$

当 $C \leqslant -1$, 改变 z 的符号后再应用同样的方法求解. 为了解当 $|C| < 1$ 时的第二个方程 (这就是 15.8 节的不可约情形), 利用相似的公式

$$\cos 3\theta = 4\cos^3\theta - 3\cos\theta,$$

得到

$$z = \cos\Big(\frac{1}{3}\arccos\ C\Big). \tag{9c}$$

在这种情形下, z 取三个值, 这是因为 $\frac{1}{3}\arccos C$ 有三个值, 它们之间相差 120° 的倍数.

习 题

1. 证明: 每个正实数有实平方根.
2. 证明: 对任意正实数 a 和任意整数 n, 方程 $x^n = a$ 有且仅有一个正实根 $\sqrt[n]{a}$.
3. 证明: 对每个 $C > -\frac{3}{8}$, $x^4 - x = C$ 有两个实根.
4. 利用正文中的 (5) 式和 $\left(\frac{\sqrt{5}}{2}\right)^2 = 1 + \frac{1}{4}$, 求 $\sqrt{5}$ 的四位小数近似值.
5. 利用 (5) 式和 $\left(\frac{5\sqrt{2}}{7}\right)^2 = 1 + \frac{1}{49}$, 求 $\sqrt{2}$ 的六位小数近似值.
6. 证明: 偶次首一多项式达到最小值 K, 并且对每个值 $C > K$, 多项式两次取值 C.
7. (a) 设 a 和 b 为正实数, $n \geqslant -1$ 为整数, 证明: 对一切充分大的正值 x, 有 $ax^{n+1} > bx^n$.
   (b) 已知多项式具有正的首项系数, 求出一个实数 M, 使得对一切 $x > M$ 有 $p(x) > 0$.
8. 证明推论 1.
9. 证明推论 2.
10. 求下列方程的实根 (取小数三位):
    (a) $3x^3 - x = \frac{1}{9}$,
    (b) $x^3 - 3x^2 + 6x = 7$,
    (c) $x^3 + 3x^2 + 2 = 0$.

*4.5 戴德金分割

想象在 x 轴上, 有理数撒布在它们各自本来的位置上. 但是, 当分割 x 轴时 (比如说, 用剪刀剪开), 我们就把全体有理数分成两类, 一类在左边, 记作 L, 一类在右边, 记作 U. 每个有理数落入这两类中的一类, 而仅当在点 $x = \frac{m}{n}$ 处分割 x 轴时,

有理数 $\frac{m}{n}$ 同时落在这两类中. 特别注意, 如果 x 在 L 中, 那么对 U 的每个 y, 有 $x \leqslant y$; 反过来, 如果对 U 中一切 y, 有 $x \leqslant y$, 那么 x 必落在 L 中. 由此引出戴德金(Dedekind)分割的想法.

一般地, 设 F 为任意有序域. 我们用 F 中的"戴德金分割"表示适合下列条件的一对非空子集 L 和 U:

(i) L 是 U 的全体元素的所有下界的集合;

(ii) U 是 L 的全体元素的所有上界的集合.

引理 1 *戴德金分割的 L 部分和 U 部分合在一起包含 F 的一切元素; 它们至多有一个公共元素.*

证明 设已知 $x \in F$, 如果对某 $a \in L$, 有 $x \leqslant a$, 那么对一切 $y \in U$, 有 $x \leqslant a \leqslant y$, 因此 $x \in L$. 否则, 由三分律, 对一切 $a \in L$, 有 $x > a$, 所以 $x \in U$. 这就证明了第一个结论: F 的每个元素不在 L 中就在 U 中. 再有, 设 a 和 b 都既在 L 中又在 U 中, 则 $a \geqslant b$(因为 $a \in U, b \in L$), 并且 $a \leqslant b$(因为 $a \in L, b \in U$), 因此 $a = b$, 这就证明了第二个结论.

如果 L 和 U 有一个公共元素 a, 则称该分割是通过a 的. 显然, 如果 L_a 是所有适合 $x \leqslant a$ 的 x 的集合, 并且 U_a 是所有适合 $x \geqslant a$ 的 x 的集合, 那么分割 (L_a, U_a) 通过 a.

戴德金分割公理 *(在有序域 F 上) 每个分割通过某元素 a.*

定理 5 *戴德金分割公理在有序域中成立当且仅当 F 是完备的有序域.*

证明 设 (L, U) 是任意分割, 如果最小上界的存在性是已知的, 则 L 具有最小上界 a. 因为 a 是 L 的上界, 所以 a 必在 U 中; 因为它是最小上界, 所以它是一切上界的下界, 因此是 U 的所有元素的下界. 根据分割定义, 这意味着 a 在 L 中, 所以给定的分割通过元素 a.

反过来, 假定戴德金分割公理成立, 并设 S 为非空有界集合. 设 U 是 S 的所有上界的集合, L 是 U 的所有下界的集合 (显然, L 包含 S). 为证明 (L, U) 是一分割, 我们只须确定 U 是 L 的所有上界的集合. 但是根据 L 的构造, U 的每个元素是 L 的上界 (对一切 $x \in L, y \in U$, 有$x \leqslant y$); 而因为 L 包含 S, 所以 U 包含 L 的一切上界. 现在根据戴德金公理, 分割线 (L, U) 通过某元素 a, 因为它是 U 的元素, 所以它是 S 的上界, 又因为它是 L 的元素, 所以它是 S 的最小上界 (即对一切 $x \in U$, 有$a \leqslant x$) . 这就完成了证明.

我们现在概述一下实数公设"实数系是完备的有序域"有关分类性质的证明.

定理 6 *任意两个完备的有序域同构.*

证明 设 F' 和 F'' 为任意两个这样的域, 根据 2.6 节定理 18 的推论 2, 它们分别包含同样的"有理数"子域 $\mathbf{Q}'$ 和 $\mathbf{Q}''$. 我们把 $\mathbf{Q}'$ 和 $\mathbf{Q}''$ 之间的同构 (保持和与积, 并

保持次序的同构) 扩张到 F' 和 F'' 之间的同构.

的确, 每个 $a' \in F'$ 定义了 F' 中的一个分割, 从而定义了 $\mathbf{Q}'$(有理数子域) 中的分割. 但根据定理 3, a' 由 $\mathbf{Q}'$ 中这个分割所确定 —— 并且 $\mathbf{Q}'$ 中每个分割 (L_R, U_R) 用这个方法确定 $a' = \text{l.u.b.}L_R = \text{g.l.b.}U_R$. $\mathbf{Q}''$ 中分割的情况类似, 因此 F' 和 F'' 的元素分别双射到 $\mathbf{Q}'$ 和 $\mathbf{Q}''$ 的分割. 这种双射显然保持次序.

最后, F' 和 F'' 中的运算可由 $\mathbf{Q}'$ 和 $\mathbf{Q}''$ 的那些运算定义, 以便把 $\mathbf{Q}'$ 和 $\mathbf{Q}''$ 的同构扩张. 更确切地说, 设 a 和 b 分别对应于 $\mathbf{Q}'$ 中分割 (L_a, U_a) 和 (L_b, U_b). 那么 $a+b$ 对应于分割 [①] (L_a+L_b, U_a+U_b), 其中 L_a+L_b 是所有和 $x+y(x \in L_a, y \in L_b)$ 的集合, 而 U_a+U_b 可类似地描述. 把正元素 a 和 b 相乘, 构成正有理数系中类似的分割. 那么 ab 对应于分割 (L_aL_b, U_aU_b), 其中 L_aL_b 是所有积 $xy(x \in L_a, y \in L_b)$ 的集合, U_aU_b 也类似地定义. 因为 $(-a)b = a(-b) = -ab$ 和 $(-a)(-b) = ab$, 因此这可扩充到一切乘积, 我们不再详细论述.

反过来, 我们可以利用分割通过整数或正整数构造实数. 我们首先证明全体有理数构成具有阿基米德性质 (定理 2 中所指出的) 的有序域. 用上段叙述的方式定义 $\mathbf{Q}$ 中分割的加法和乘法, 我们可以证明 $\mathbf{Q}$ 中分割构成满足戴德金分割公理的有序域, 因而给出完备的有序域. 但是证明很长, 会把我们引入迷途, 因此我们只是叙述一下结果.

定理 7 *有一个且仅有一个 (同构的域除外) 完备的有序域.*

不用戴德金分割, 而通过有理数, 把实数看作有理数序列的极限, 也可以构造实数. [②]

习 题

1. 证明: 如果(L, U)和(L', U')是有理数域中的分割, 那么每个有理数 (至多有一个例外) 都能表为 $x+y(x \in L, y \in L')$, 或者表为 $u+v(u \in U, v \in U')$.
2. 叙述并证明对于正有理数在乘法下的类似于习题 1 的一个定理.
3. 为什么这个定理对于负有理数不成立?
4. 证明: 对于每个 $\varepsilon > 0$, 存在充分大的 n 使得 $10^{-n} < \varepsilon$.
5. 有序域 F 中的戴德金分割有时定义为 F 的一对子集 L' 和 U', 它们适合: F 的每个元素或者在 L' 中或者在 U' 中, 并且当 $x \in L', y \in U'$ 时, 有 $x < y$. 通过添加或删去相应的单个元素, 可以证明: 这种类型的每个分割给出正文中定义的分割 (L, U), 反之亦然.

---

① 在某些情况下, (L_a+L_b, U_a+U_b) 不会是分割, 因为数 $a+b$ 不在 L 中也不在 U 中; 但是, 如果把遗漏的数添到 L 和 U 上去, 我们便得到分割. 类似的情况适合于下面的 L_aL_b.

② 参看 C. C. MacDuffee, *Introduction to Abstract Algebra* (New York, Wiley, 1940) 的第 VI 章的论述.

6. 设 t 为有序整环 D 中的元素, $0<t<1$, 证明：$s=2-t$ 具有性质 $s>1, st\leqslant 1$.
7. 设 D 为不同构于 $\mathbf{Z}$ 的完备的有序整环. 证明：D 包含适合 $0<t<1$ 的元素 t. 设 b 和 c 为 D 的任意正元素, 证明：对某整数 n, $t^n b<c$.

*8. 利用习题 6 和习题 7 证明：任意完备的有序整环或者同构于 $\mathbf{Z}$, 或者同构于 $\mathbf{R}$. (提示：求出 $b>1$ 的逆元素, 考虑满足 $xb\leqslant 1$ 的所有的 x.)

9. (a) 证明：$\mathbf{R}$ 的任意自同构保持关系 $x\leqslant y$. (提示：$x\leqslant y$ 当且仅当 $z^2=y-x$ 有解.)
   (b) 利用 (a) 证明: $\mathbf{R}$ 唯一的自同构是平凡同构 $x\mapsto x$.

*10. 证明：如果 $D=F$ 为有序域, 并且对每个有理函数

$$R(x)=\frac{b_0+b_1x+\cdots+b_rx^r}{a_0+a_1x+\cdots+a_nx^n}\neq 0,\quad a_nb_r\neq 0,$$

我们规定 $R(x)>0$ 的意思是 $a_nb_r>0$, 那么 $F(x)$ 成为有序域.

*11. 证明：在习题 10 中, $R(x)>0$ 当且仅当对 F 中一切充分大的 t, 有 $R(t)>0$.

第5章 复　　数

5.1 复数的定义

如果把实数系 $\mathbf{R}$ 扩张成较大的复数域 $\mathbf{C}$, 那么在解析函数论和微分方程论中, 特别是在代数学中, 很多代数定理的描述将更为简洁. 我们现在就来定义复数域, 并且指出, 如果我们想要使每个多项式方程都有根, 由实数域扩张而得到的域就是复数域.

定义 *复数就是实数偶 (x,y)——x 称为 (x,y) 的实分量, y 称为 (x,y) 的虚分量. 根据法则*

$$(x,y)+(x',y')=(x+x',y+y'), \tag{1}$$

$$(x,y)\cdot(x',y')=(xx'-yy',xy'+yx'), \tag{2}$$

施行复数相加和相乘. 这样定义的复数系记作 $\mathbf{C}$.

我们认为上面的定义不是靠神灵的力量, 而是通过简单的代数运算而得到的. 首先, 观察出方程 $x^2=-1$ 没有实根 (x^2 决不能是负数). 这就暗示我们要引进一个虚数 i, 它满足 $\mathrm{i}^2=-1$, 此外还满足普通的代数定律. 用确切的话来说, 它提出一个表面上讲得通的假设: 存在一个包含元素 i 同时包含实数域 $\mathbf{R}$ 的整环 D.

在 D 中, 任意形为 $x+y\mathrm{i}$(x,y 为实数) 的表达式将表示一个元素. 此外, 由整环的定义 (普通的代数定律) 有

$$(x+y\mathrm{i})\pm(x'+y'\mathrm{i})=(x\pm x')+(y\pm y')\mathrm{i}, \tag{1$'$}$$

$$(x+y\mathrm{i})\cdot(x'+y'\mathrm{i})=xx'+(xy'+yx')\mathrm{i}+yy'\mathrm{i}^2. \tag{2$'$}$$

因为 $\mathrm{i}^2=-1$, 由 (2$'$) 得

$$(x+y\mathrm{i})\cdot(x'+y'\mathrm{i})=(xx'-yy')+(xy'+yx')\mathrm{i}. \tag{2$''$}$$

于是得到一个推论是, 由 $\mathbf{R}$ 和 i 生成的 D 的子整环包含一切形为 $x+y\mathrm{i}$ 的元素, 而不包含其他任何元素.

再有, 由 $x+y\mathrm{i}=x'+y'\mathrm{i}$ 推出 $x-x'=(y'-y)\mathrm{i}$, 因此两边平方得 $(x-x')^2=-(y'-y)^2$, 又因 $(x-x')^2\geqslant 0, -(y'-y)^2\leqslant 0$, 除非 $x=x', y=y'$, 否则上面等式不成立. 总之, 不同的实数偶 (x,y) 确定 D 中不同的元素 $x+y\mathrm{i}$. 这就建立了, $\mathbf{C}$ 中元

素和由 $\mathbf{R}$ 与 i 生成的 D 的子整环中元素之间的一一对应 $(x,y)\leftrightarrow x+y\mathrm{i}$. 比较公式 $(1')\sim(2'')$ 和 $(1)\sim(2)$, 我们看到这种对应保持和与积, 因此是一个同构. 这就证明了

定理 1 设 D 为包含实数系 $\mathbf{R}$ 和 -1 的平方根 i 的任意整环. 那么由 $\mathbf{R}$ 和 i 生成的 D 的子整环与 $\mathbf{C}$ 同构.

我们现在证明我们的猜想, 确定存在一个整环 D, 它包含全体实数和 -1 的平方根.

定理 2 按上面定义的复数系是一个域, 它包含一个与 $\mathbf{R}$ 同构的子域, 并包含方程 $x^2+1=0$ 的根.

证明 对于实数偶 (x,y), 加法的交换律和结合律成立, $(0,0)$ 是加法单位元素, $(-x,-y)$ 是 (x,y) 的加法逆元素, 这些都是下述事实的直接结论: 数偶的实分量和虚分量是独立相加的, 而相应的定律对于它们都是成立的.

类似地, 乘法的交换律和结合律成立, $(1,0)$ 是乘法单位元素, 每个 $(x,y)\neq(0,0)$ 都有乘法逆元素

$$(x,y)^{-1}=\left(\frac{x}{x^2+y^2},\frac{-y}{x^2+y^2}\right). \tag{3}$$

这可由 5.2 节中建立的事实得到, 在 5.2 节中指出, 几个复数相乘时, 它们的"辐角"和"绝对值"是分开来进行运算的, 这些运算本身满足交换律和结合律. 可是在这里, 检验这些定律所采用的办法是直接代入定义 (2), 只有结合律的计算比较冗长, 我们略去了它们的详细验证.

最后, 类似地直接代入定义, 我们可以验证分配律. 设 $z=(x,y), z'=(x',y'), z''=(x'',y'')$. 那么代入定义 (1) 和 (2), 有

$$\begin{aligned}
z(z'+z'') &= (x,y)(x'+x'',y'+y'')\\
&= (x(x'+x'')-y(y'+y''),x(y'+y'')+y(x'+x'')),\\
zz'+zz'' &= (xx'-yy',xy'+yx')+(xx''-yy'',xy''+yx'')\\
&= (xx'-yy'+xx''-yy'',xy'+yx'+xy''+yx'').
\end{aligned}$$

从这两个表达式可以直接验证 $z(z'+z'')=zz'+zz''$.

在这个由数偶组成的域 $\mathbf{C}$ 中, 我们利用定理 1 中所用过的对应 $(x,y)\leftrightarrow x+y\mathrm{i}$, 可以在 $\mathbf{C}$ 中找到一个实数子域. 按照这种对应, 实数 x 对应于第二项为零的数偶, (0, 1) 对应于 i. 特别是, 如果定义 (1) 和 (2) 中的第二个分量 y 和 y' 都是零时, 那么第一个分量 x 和 x' 的相加和相乘恰好与实数 x 和 x' 的相加和相乘一样. 这正是我们所要认识的: 对应 $x\leftrightarrow(x,0)$ 是实数域 $\mathbf{R}$ 到 $\mathbf{C}$ 的子域的一个同构. 在上面这种情况下, 我们认为, 每个这样特殊的复数 $(x,0)$ 只与相应的实数 x 等同.

最后, 我们希望 -1 的平方根相当于数偶 (0, 1). 事实上, 定义 (2) 的特殊情况表明, $(0,1)^2=(-1,0)=-1$. 因此我们定义i 是数偶 (0, 1). 那么任意数偶 (x,y) 具

有形式

$$(x,y)=(x,0)+(0,y)=(x,0)+(y,0)(0,1)=x+y\mathrm{i}. \tag{4}$$

记号 $x+y\mathrm{i}$ 是很方便的, 后面我们都用它来代替 (x,y). 为简洁起见, 我们还常常写作 $z=(x,y)=x+y\mathrm{i}, w=(u,v)=u+v\mathrm{i}, c=(a,b)=a+b\mathrm{i}$, 等等 —— 换句话说, 我们用单个字母表示复数, 用字母表中紧靠着它的前两个字母分别表示这个复数的实分量和虚分量.

习 题

1. 验证复数乘法满足交换律和结合律.
2. 用公式 (3) 验证 $(x,y)(x,y)^{-1}=(1,0)$.
3. 解方程 $(1,1)(x,y)=(2,1)$,
   (a) 化为一对关于变量 x,y 的联立线性方程,
   (b) 用公式 (3).
4. 分别求出满足下列关系的复数 $z=x+y\mathrm{i}$ 和 $w=u+v\mathrm{i}$:
   (a) $z+\mathrm{i}w=1, \mathrm{i}z+w=1+\mathrm{i}$,
   (b) $(1+\mathrm{i})z-\mathrm{i}w=3+\mathrm{i}, \quad (2+\mathrm{i})z+(2-\mathrm{i})w=2\mathrm{i}$.
5. 求出方程 $z^2=-a$(a 为任意正实数) 的全部复根, 并验证你的答案.
6. 描述由 i 和全体有理数生成的 $\mathbf{C}$ 的子域.
7. 如果 D 是交换环, 定理 1 还成立吗? 给出详细证明.
8. (a) 证明: $z^2=a+b\mathrm{i}$ 有解 $x+y\mathrm{i}$, 其中 $x=\left[\dfrac{a+\sqrt{a^2+b^2}}{2}\right]^{\frac{1}{2}}, \quad y=\dfrac{b}{2x}$.
   (b) 证明: 方程的解还可表成
   $$y=\left[\frac{\sqrt{a^2+b^2}-a}{2}\right]^{\frac{1}{2}}, \quad x=\frac{b}{2y}.$$
   (注意, 当 a 是负数, 并且 $\dfrac{b}{a}$ 很小时, 这组公式在数值计算中更为精确.)
9. 方程 $z^3+3\mathrm{i}z=3+\mathrm{i}$ 有一个根 $-\mathrm{i}$, 计算另一个根, 并表示成小数形式.

*10. 证明: 如果 F 为任意有序域, 那么存在一个比它大的域 F^*, 包含着与 F 同构的子域和 -1 的平方根.

*11. 用定理 1 和定理 2 的方法, 不借助于实数, 证明: 有理数域 $\mathbf{Q}$ 可以扩张到较大的域 $\mathbf{Q}(\sqrt{2})$, 该域包含 $\mathbf{Q}$ 和 2 的平方根.

12. 证明: 不可能存在 "正复数" 的定义, 使 $\mathbf{C}$ 构成有序域.

5.2 复 平 面

全体复数到笛卡儿平面的全体点上有一个基本的一一映射. 即每个复数 $z=$

$x + y\mathrm{i}$ 映射到点 $P = (x, y)$ 上, 这个点以 z 的实分量 x 为横坐标, 以虚分量 y 为纵坐标.

极坐标可以用在这个平面上. 回想平面上的点 P, 可以认为每个复数是由两个极坐标 r 和 θ 唯一确定的, 这里 r 是联结点 P 到原点的线段 $\overline{Oz}$ 的长度 (非负的), 而 θ 是由 x 轴到线段 $\overline{Oz}$ 的夹角 (图 5-1), 所以

$$|z| = r = (x^2 + y^2)^{\frac{1}{2}}, \arg z = \theta = \arctan\frac{y}{x}. \qquad (5)$$

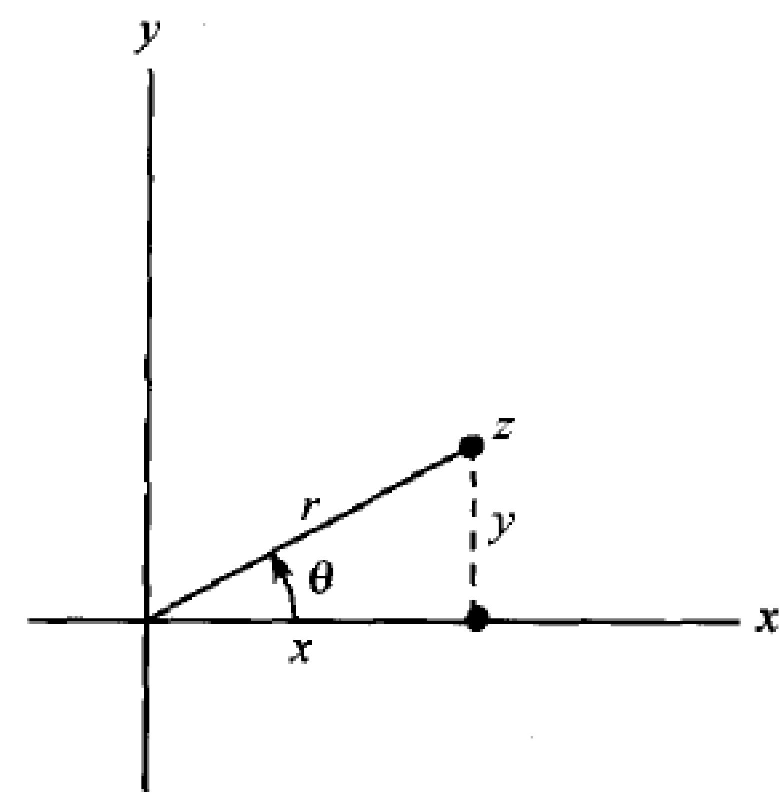


图 5-1

我们把 r 称为复数 z 的绝对值, 把 θ 称为 z 的辐角. r 和 θ 按下式确定 x 和 y

$$x = r\cos\theta, \quad y = r\sin\theta, \quad z = r(\cos\theta + \mathrm{i}\ \sin\theta). \qquad (6)$$

这就是通常由极坐标到直角坐标的变换公式. 我们还可以把 (6) 式写成 $z = r\mathrm{e}^{\mathrm{i}\theta}$ 的形式, 这是因为, 由通常的泰勒级数展开式得到

$$\mathrm{e}^{\mathrm{i}\theta} = 1 + \mathrm{i}\theta + \frac{(-1)\theta^2}{2!} + \frac{(-\mathrm{i})\theta^2}{3!} + \cdots = \cos\theta + \mathrm{i}\ \sin\theta.$$

绝对值和辐角的重要性主要反映在隶莫弗 (De Moivre) 公式, 这个公式叙述如下:

定理 3 复数乘积的绝对值等于因子的绝对值之积, 乘积的辐角等于因子的辐角之和, 换句话说,

$$|zz'| = |z| \cdot |z'|, \quad \arg zz' = \arg z + \arg z'. \qquad (7)$$

证明 因为按 (6) 式, 有 $z = r(\cos\theta + \mathrm{i}\ \sin\theta), z' = r'(\cos\theta' + \mathrm{i}\ \sin\theta')$, 代入定义 (2) 中我们得到

$$zz' = rr'[(\cos\theta\cos\theta' - \sin\theta\sin\theta') + \mathrm{i}(\cos\theta\sin\theta' + \sin\theta\cos\theta')];$$

由熟知的三角公式, 这就是

$$zz' = rr'[\cos(\theta + \theta') + \mathrm{i}\ \sin(\theta + \theta')].$$

这就给出了结论 (7).

复数绝对值的乘法性质与加法性质 (不等式) 其表达形式同实数一样, 即

$$|z| > 0 \quad 除非 \quad z = 0, |0| = 0; \qquad (8)$$

$$|z+z'| \leqslant |z|+|z'|. \tag{9}$$

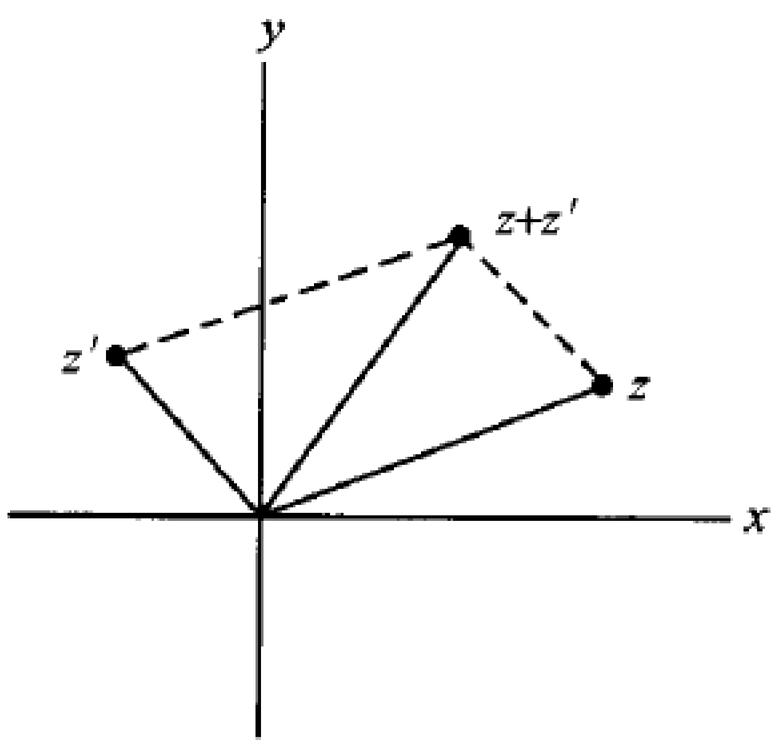


图 5-2

为了证明这些，注意公式 (1) 意味着 $z+z'$ 可以通过画以 z,0 和 z' 为三个顶点的平行四边形 (图 5-2) 来求得，第四个顶点就是 $z+z'$. 由于复数的绝对值等于相应线段的几何长度，现在可以推出公式 (8) 和公式 (9).

复 n 次单位根可以用三角方法求得. 从隶莫弗从式 (7) 直接得出

$$[r(\cos\theta+\mathrm{i}\ \sin\theta)]^{-1}=\frac{1}{r}[\cos(-\theta)+\mathrm{i}\ \sin(-\theta)].$$

进一步得到，$z^n=1$ 当且仅当 $|z|^n=1$，并且 $n\cdot\arg z$ 是 2π 的整数倍 $2k\pi$. 因为 $|z|\geqslant 0$，所以 $|z|=1$. 因为 $\arg z$ 在 $0\leqslant\theta<2\pi$ 上是单值的，所以 $z^n=1$ 确有 n 个解. 在直角坐标中，它们是

$$1, \cos\frac{2\pi}{n}+\mathrm{i}\ \sin\frac{2\pi}{n}, \cdots, \cos\frac{2\pi(n-1)}{n}+\mathrm{i}\ \sin\frac{2\pi(n-1)}{n}.$$

如果我们用 ω 表示 $\cos\frac{2\pi}{n}+\mathrm{i}\ \sin\frac{2\pi}{n}$，则可得到这些 n 次单位根的另一种表示：1, $\omega, \omega^2, \cdots, \omega^{n-1}$. 用几何语言叙述就是

定理 4 全体复 n 次单位根是单位圆 $|z|=1$ 内接正 n 边形的 n 个顶点.

更一般地，考虑方程 $z^n=c$，其中 $c\neq 0$ 为任意复数. 在极坐标中，方程的一个解是

$$z_0=|c|^{\frac{1}{n}}(\cos\theta+\mathrm{i}\ \sin\theta) \quad 其中 \quad \theta=\frac{1}{n}\arg c.$$

此外，wz_0 是 $x^n=c$ 的根当且仅当 $c=(wz_0)^n=w^nz_0^n=w^nc$，因此 $w^n=1$. 于是 c 的 n 次根是 $z_0, \omega z_0, \omega^2 z_0, \cdots, \omega^{n-1}z_0$，这里 ω 是上面定义的复 n 次单位根. 特别地，它们也可以用正多边形的 n 个顶点表示.

对 $c=a+b\mathrm{i}$ 的 n 次根 $z_0, \omega z_0, \omega^2 z_0, \cdots, \omega^{n-1}z_0$，我们可以借助于三角函数表和对数表，很容易地计算出它们的数值，从恒等式

$$\ln|z_0|=\ln|c|^{\frac{1}{n}}=\frac{1}{n}\ln(a^2+b^2)^{\frac{1}{2}}=\frac{1}{2n}\ln(a^2+b^2)$$

出发，我们可以计算 $|z_0|$. 根据隶莫弗公式 (7)，有 $\arg z_0=\frac{1}{n}\arctan\frac{b}{a}$ 和 $\arg \omega^k z_0=\frac{1}{n}\arctan\frac{b}{a}+\frac{360k}{n}$. 这里的单位是度. 最后由公式

$$z=r(\cos\theta+\mathrm{i}\ \sin\theta)=|z|\cos(\arg z)+\mathrm{i}|z|\sin(\arg z)$$

完成计算.

每个复 n 次单位根 ω 满足一个有理数域上不可约的有理系数多项式方程. 这些方程称为“分圆”方程, 在方程式理论中起着重要的作用.

由定义, 每个 n 次单位根满足方程 $z^n-1=0$. 此外, 除了 $z=1$ 的其他所有根满足

$$q_n(z)=\frac{z^n-1}{z-1}=z^{n-1}+z^{n-2}+\cdots+z+1=0. \tag{10}$$

在 3.10 节中用爱森斯坦判别准则证明了当 $n=p$ 是素数时, $q_p(z)$ 是不可约的.

如果 n 不是素数, 那么情况就复杂了. 例如, 当 $n=4, z^3+z^2+z+1=(z+1)(z^2+1)$ 是可约的. 一般地, 我们可以从 (10) 中分解出 k 次单位根所满足的分圆多项式, 这里 k 取遍 n 的所有真因子. n 次单位根, 如果对所有的 $k<n$, 它不是 k 次单位根, 则称它为 *n 次本原单位根*. (例如, 四次本原单位根是 i 和 $-$i.) n 次本原单位根是 ω^m, 其中 m 与 n 互素, 它们都满足有理数域上同一个不可约方程. 但是这一结论的证明和这个方程次数的计算, 需要更多的数论知识.

习　题

1. 用棣莫弗公式证明复数乘法的交换律和结合律以及乘法逆元素的存在性.
2. 描述对应关系 $z\longmapsto z\mathrm{i}$ 的几何意义.
3. 求三次单位根和五次单位根的实分量与虚分量, 计算到小数点后四位 (用三角函数表).
4. 求 $2+2\mathrm{i}$ 的立方根和四次根, 计算到小数点后四位.
5. 列出全部 12 次本原单位根, 并在图纸上把它们画在一个大单位圆上.
6. 用几何语言描述变换 $z\longmapsto cz+d(c,d\in\mathbf{C},c\neq 0)$ 的效果. 当 $|c|=1$ 时是什么情况? (*提示*: 使用“平移”、“旋转”和“放大”等词.)
7. 找出 z^6-1 在有理数域 $\mathbf{Q}$ 上的不可约因子.
8. (a) 证明: $\omega=\cos\frac{2\pi}{n}+\mathrm{i}\sin\frac{2\pi}{n}$ 是 n 次本原单位根.
   (b) 证明: ω^m 是 n 次本原单位根当且仅当 m 与 n 互素.

5.3 代数基本定理

我们在 5.1 节中看到, 实数系 $\mathbf{R}$ 添加方程 $z^2+1=0$ 的一个虚根 i 就得到复数系. 但是为什么就到此为止了呢? 为什么不打算添加其他多项式方程的“虚”根以便得到更大的域呢? 所谓代数基本定理就回答了这个问题: 一旦添加上 i, 那么每个多项式方程就必有 (复) 根, 所以为解方程我们就不需要再选另外的虚根.

定理 5 (欧拉–高斯) *每个正次数的复系数多项式必有复根.*

已经知道这个著名定理的很多证明方法. ① 所有的证明都包含着像第 4 章引进的那些非代数概念, 这里我们选择了一个证明, 它的非代数部分在直观上好像特别容易说明. 我们不从第 4 章有关的公理来详细证明非代数部分.

证明 因为多项式 $p(z)=a_mz^m+a_{m-1}z^{m-1}+\cdots+a_0$, 其中 $a_m\neq 0$ 与多项式

$$q(z)=z^m+\frac{a_{m-1}}{a_m}z^{m-1}+\cdots+\frac{a_0}{a_m}=z^m+c_{m-1}z^{m-1}+\cdots+c_0$$

有相同的根, 所以只须讨论首项系数为 1 的情况.

这种情况下, 我们画两个复平面, 一个标记为"z 平面"另一个标记为"w 平面". 已知函数 $q(z)$ 把 z 平面的每个点 $z_0=(x_0,y_0)$ 映射到 w 平面的点 $w_0=q(z_0)$ 上. 此外, 如果 z 描绘 z 平面上的一条连续曲线, 那么 $q(z)$(是可微的) 将描绘 w 平面上的一条连续曲线, 我们的目的是证明, w 平面的原点 0 是 z 平面上某个点 z 的"像" $q(z)$, 或者与之同样的是证明 z 平面上某个圆的像通过 w 平面的原点 0.

对每个固定的 $r>0$, 函数 $w=q(re^{i\theta})$ 确定了 w 平面上的一条闭曲线 γ_r', 即 z 平面上以原点 0 为中心、r 为半径的圆 $\gamma_r:|z|=r(z=re^{i\theta})$ 的像. 对每个固定的 r, 考虑线积分 ②

$$\phi(r,\theta)=\int_0^\theta \mathrm{d}(\arg w)=\int_0^\theta \frac{u\mathrm{d}v-v\mathrm{d}u}{u^2+v^2},$$

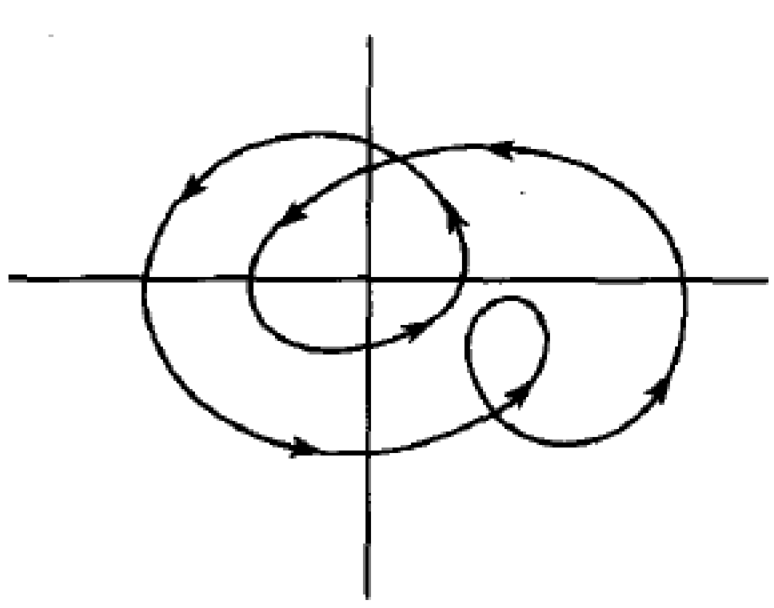


图 5-3

这是对任何不通过原点 $w=0$ 的曲线 γ_r' 来定义的. (如果 γ_r' 通过 $w=0$, 那么定理 5 的结论就立即可得.) 几何上显然有 $\phi(r,2\pi)=2\pi n(r)$, 这里分支数$n(r)$ 是曲线 γ_r' 绕原点反时针旋转的次数. 例如, 图 5-3 描绘的曲线, 其 $n(r)=2$.

现在考虑 $n(r)$ 随 r 的变化情况. 因为 $q(re^{i\theta})$ 是连续函数, 所以除了 γ_r' 通过原点外, $n(r)$ 也是随 r 连续变化的. 再有, $n(0)=0$(除非 $c_0=0$, 这时 0 是方程的根). 现在假定 $c_0\neq 0$, 我们将证明, 当 r 充分大时, $n(r)$ 就是 $q(z)$ 的次数m. 实际上, 设

① 例如参见, L. E. Dickson, *New First Course in the Theory of Equations* (New York: Wiley, 1939), 附录, 或 L. Weisner, *Introduction to the Theory of Equations* (New York: Macmillan, 1938), p. 145.

② 在证明线积分的存在时, 必须用到 $\mathbf{R}$ 的完备性. 恒等式

$$\mathrm{d}(\arg w)=\frac{u\mathrm{d}v-v\mathrm{d}u}{u^2+v^2}$$

成立是因为 $\arg w=\arctan\frac{v}{u}$.

$$q(z) = z^m + c_{m-1}z^{m-1} + \cdots + c_1 z + c_0 = z^m\Big(1 + \sum_{k=1}^{m} c_{m-k}z^{-k}\Big).$$

根据隶莫弗公式 (7), 有

$$\arg q(z) = m \arg z + \arg\Big(1 + \sum_{k=1}^{m} c_{m-k}z^{-k}\Big).$$

因此, 当 z 沿着圆 γ_r 作反时针方向变化一周时, $\arg q(z)$ 得到的改变量是 $\arg z$ 的改变量的 m 倍 (即 $m \cdot 2\pi$) 加上 $\arg\Big(1 + \sum_{k=1}^{m} c_{m-k}z^{-k}\Big)$ 的改变量. 可是当 $|z| = r$ 充分大时, 由公式 (8) 和公式 (9) 可知

$$1 + \sum_{k=1}^{m} c_{m-k}z^{-k} = u$$

停留在圆 $|u-1| < \frac{1}{2}$ 中, 因此绕原点只转了零次(画个图加以解释).

我们得出结论: 当 r 充分大时, $n(r) = m, \arg q(z)$ 的总改变量是 $2\pi m$. 但是当 r 变化时, γ_r' 是连续变形的 (因为 $q(z)$ 是连续的). 然而, 几何上显然有,[①] 一条绕原点 $m(\neq 0)$ 次的曲线, 如果不是它变形的某一步通过原点, 这条曲线就不能连续地变形成一点. 由此推出, 对某个 r, γ_r' 必通过原点, 这就出现 $q(z) = 0$! 证毕

作为推论, 我们注意, 如果 $p(z_1) = 0$, 那么根据余数定理 (3.5 节), 我们可以写成 $p(z) = (z - z_1)r(z)$. 如果 $p(z)$ 的次数为 $m, m > 1$, 则商式 $r(z)$ 具有正次数, 因此它也有一个复根 $z = z_2$. 如此进行下去, 我们就找到 $p(z)$ 的 m 个线性因子, 如

$$p(z) = c(z - z_1)(z - z_2)\cdots(z - z_m). \tag{11}$$

由此得到, *$\mathbf{C}$上的不可约多项式只能是线性的*. 这个推论和第 3 章的唯一因子分解定理合在一起得出

定理 6 *任意复系数多项式可按一种且仅按一种方式写成 (11) 的形式.*

在 (11) 中 $p(z)$ 的根显然是 $z_1, \cdots, z_m$, 这是因为乘积为零当且仅当它其中一个因子为零. 如果因子 $(z - z_i)$ 重复出现, 那么它重复出现的次数称为根 z_i 的*重数*. 在微积分学中, 这个可以定义为 $p(z)$ 在 z_i 点为零的 "阶数": 使得 $p(z)$ 和它的前 $\nu - 1$ 阶导数在 z_i 点都为零的最大整数 ν.

习 题

1. 不用 3.8 节一般的唯一性定理, 证明: 分解式 (11) 的唯一性.

① 这作为平面拓扑学中的一个定理已经证明了. 例如参看 S. Lefschetz, *Introduction to Topology* (Princeton University Press, 1949), p. 127.

2. 证明：任何有理复函数, 如果对所有的 z 都取有限值, 则它是多项式.
3. 所有复数偶 (w, z), 当相加和相乘遵循法则 (1) 和 (2) 时, 它构成含有单位元素的交换环吗？构成域吗？
4. 证明：任意二次多项式可以通过 $\mathbf{C}[z]$ 的适当的自同构得到形式 $cz(z-1)$ 或 cz^2.
5. (a) 用马克劳林 (Maclaurin) 级数证明公式 $\mathrm{e}^{\mathrm{i}x} = \cos x + \mathrm{i}\sin x$.
   (b) 证明每个复数可以写成 $r\mathrm{e}^{i\theta}$.
   (c) 推导恒等式
$$\cos z = \frac{\mathrm{e}^{\mathrm{i}z}+\mathrm{e}^{-\mathrm{i}z}}{2}, \quad \sin z = \frac{\mathrm{e}^{\mathrm{i}z}-\mathrm{e}^{-\mathrm{i}z}}{2\mathrm{i}}.$$
6. 利用部分分式证明：域 $\mathbf{C}$ 上的任意有理函数可以写成一个多项式加上一些有理函数之和, 这些有理函数的分子是常数, 分母是线性函数的幂.
7. 分解 $z^2 + z + 1 + \mathrm{i}$.

5.4 共轭数与实多项式

在复数域 $\mathbf{C}$ 上, 方程 $z^2 = -1$ 有两个根 i 和 $-\mathrm{i} = 0 + (-1)\mathrm{i}$. 对应 $x + y\mathrm{i} \longmapsto x + y(-\mathrm{i}) = x - y\mathrm{i}$ 把第一个根映射到第二个根, 反过来把第二个根映射到第一个根, 而它们保持所有实数不变. 而且这个对应把和映射到和, 把积映射到积, 这可以通过直接代入公式 (1) 和公式 (2) 或者应用定理 1 来验证. 换句话说, 这个对应是 $\mathbf{C}$ 的一个自同构($\mathbf{C}$ 到自身的同构).

我们可以更简洁地把这个对应叙述如下. 把数 $x - y\mathrm{i}$ 称为复数 $z = x + y\mathrm{i}$ 的“共轭” z^*①. 对应 $z \longmapsto z^*$ 是 $\mathbf{C}$ 上周期为 2 的自同构, 这因为

$$(z_1 + z_2)^* = z_1^* + z_2^*, \quad (z_1 z_2)^* = z_1^* z_2^*, \quad (z^*)^* = z. \tag{12}$$

在几何上这个对应相当于复平面关于 x 轴的一个反射；与其共轭相等的数只有实数.

共轭复数在数学中和物理学中 (特别在波动力学中) 是很有用的. 在使用它们的时候, 记住下面一些简单公式是方便的：

$$|z|^2 = zz^*, \qquad z^{-1} = \frac{z^*}{|z|^2}.$$

用这些公式可使我们从定理 6 很容易地推出实系数多项式的分解定理.

引理 *实系数多项式的非实复根是以一对共轭复数的形式出现的.*

这推广了下面熟知的事实：二次多项式 $ax^2 + bx + c$, 当判别式 $b^2 - 4ac < 0$ 时, 有两个共轭复根 $x = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$.

① 现在通常写作 $\bar{z}$.

证明 设 $p(z)$ 为已知多项式, 我们可以把它写成 (11) 的形式, 其中 z_i 是复数 (不是通常的实数). 因为作用在这些根 z_i 上的一个对应 $z_i \longmapsto z_i^*$ 是自同构, 所以它把 $p(z)$ 映射到另一个多项式

$$p^*(z) = c^*(z - z_1^*)(z - z_2^*) \cdots (z - z_n^*).$$

这个多项式的每个系数是 $p(z)$ 中相应系数的共轭. 但因 $p(z)$ 的全体系数都是实数, 所以 $p(z) = p^*(z)$. 因此分解式 (11) 是唯一的, $c = c^*$ 是实数, 并且 z_i 是实数或者是成对出现的共轭复数.

定理 7 *任意实系数多项式可以分解成 (实) 线性多项式和判别式为负的 (实) 二次多项式.*

证明 上面引理中的实根 z_i 给出 (实) 线性因子 $(z - z_i)$. 一对共轭复根 $a + b\mathrm{i}$ 和 $a - b\mathrm{i}(b \neq 0)$ 可以合起来有

$$[z - (a + b\mathrm{i})][z - (a - b\mathrm{i})] = z^2 - 2az + (a^2 + b^2),$$

它给出 $p(z)$ 的一个实系数二因子, 其判别式为

$$4a^2 - 4(a^2 + b^2) = -4b^2 < 0.$$ 证毕

反过来, 线性多项式和判别式为负的二次多项式在实数域上是不可约的 (后者是因为它们只有复数根, 因此没有线性因子). 定理 7 所描述的因子分解是唯一的, 这可作为一个推论.

习　题

1. 解方程:

   (a) $(1 + \mathrm{i})z + 3\mathrm{i}z^* = 2 + \mathrm{i}$,　(b) $zz^* + 2z = 3 + \mathrm{i}$,　(c) $zz^* + 3(z - z^*) = 4 - 3\mathrm{i}$.

2. 解方程:

   (a) $zz^* + 3(z + z^*) = 7$,　(b) $zz^* + 3(z + z^*) = 3\mathrm{i}$.

3. 解联立方程:

$$\begin{cases} \mathrm{i}z + (1 + \mathrm{i})w = 3 + \mathrm{i}, \\ (1 + \mathrm{i})z^* - (6 + \mathrm{i})w^* = 4. \end{cases}$$

4. 给出 4.4 节定理 4 推论 2 的独立的证明.
5. 证明: 如果我们在实数系上添加一个任意非线性不可约的实系数多项式的虚根, 则可得到一个与 **C** 同构的域.
6. 证明: 在任意有序域上, 如果 $b^2 - 4ac < 0$, 则 $ax^2 + bx + c$ 是不可约的.

7. 证明：保持所有实数都不变的 $\mathbf{C}$ 的每个自同构或者是恒等自同构 ($z \longmapsto z$)，或者是自同构 $z \longmapsto z^*$.

*5.5 二次方程与三次方程

5.3 节中我们证明了任意复系数多项式根的存在性，但没有指出如何有效地把根计算出来. 在 5.5 节和 5.6 节中，我们将指出如何计算二次方程、三次方程和四次方程的根. 计算过程中只包含四种有理运算 (加、减、乘、除) 和开 n 次方根运算. 5.1 节和 5.2 节中我们已指出如何进行复数的这些运算. 下面讲的计算过程也可用于任何别的域上，在这些域上，任意元素的 n 次根是可以构造的，而且 $1+1\neq 0, 1+1+1\neq 0$.

二次方程可以用中学代数的“配方”方法求解. 方程

$$az^2+bz+c=0 \quad (a\neq 0), \tag{13}$$

等价于 (具有同样的根) 较简单的方程

$$z^2+Bz+C=0 \quad \left(B=\frac{b}{a}, C=\frac{c}{a}\right). \tag{14}$$

如果令 $w=z+\dfrac{B}{2}\left(\text{即} z=w-\dfrac{B}{2}\right)$，以便配成完全平方，我们可以看到 (14) 式等价于

$$w^2=\frac{B^2}{4}-C. \tag{15}$$

对 w, B, C 代回 z, a, b, c，这就得出

$$z=w-\frac{B}{2}=\frac{-b+\sqrt{b^2-4ac}}{2a}, \tag{16}$$

根据 5.2 节，所以有两个解.

三次方程可以类似地求解. 首先像 4.4 节那样，把三次方程化为形式

$$z^3+pz+q=0, \tag{17}$$

然后做维特 (Vieta) 变换 $z=w-\dfrac{p}{3w}$，结果得到 (有些项已消去)

$$w^3-\frac{p^3}{27w^3}+q=0. \tag{18}$$

用 w^3 乘以各项，我们得到关于 w^3 的二次方程. 这个方程可以根据 (16) 式求解，得出

$$w^3=-\frac{q}{2}+\sqrt{\frac{q^2}{4}+\frac{p^3}{27}} \quad (\text{两个值}). \tag{19}$$

这给出 w 的 6 个三次根形式的解. 把这些解代入公式 $z = w - \frac{p}{3w}$, 我们就得到 z 的三对解, 成对的两个解是相等的.

阐述一下前面定理 6 中的公式是有趣的. 例如, 二次多项式的情形, 记

$$z^2 + Bz + C = (z - z_1)(z - z_2),$$

我们有

$$z_1 + z_2 = -B, z_1z_2 = C, \quad 因而 (z_1 - z_2)^2 = B^2 - 4C. \tag{20}$$

$B^2 - 4C = D$ 这个量是 (14) 的判别式. 用原来 (13) 式中的系数表示, $D = \frac{b^2 - 4ac}{a^2}$.

类似地, 设 z_1, z_2, z_3 是简化的三次方程 (17) 的根, 则

$$z_1 + z_2 + z_3 = 0, \quad z_1z_2 + z_2z_3 + z_3z_1 = p, \quad z_1z_2z_3 = -q. \tag{21}$$

合并前两个关系式, 我们得到公式

$$p = z_1z_2 - z_3^2, \quad (z_1 - z_2)^2 = -4p - 3z_3^2, \quad z_1^2 + z_2^2 + z_3^2 = -2p. \tag{22}$$

我们现在用

$$D = \prod_{i<j}(z_i - z_j)^2 = P^2, \quad 其中 P = (z_1 - z_2)(z_2 - z_3)(z_1 - z_3) \tag{23}$$

来定义三次方程的判别式. 把 P 平方, 再利用 (22) 式的第二个关系式, 通过一些计算后我们就得到

$$D = -4p^3 - 27q^2, \tag{24}$$

它可以用来简化 (19), 得 $w = -\frac{q}{2} + \frac{\sqrt{-D}}{6}$.

定理 8 实系数二次方程或三次方程, 如果它的判别式非负, 则它有实根; 如果它的判别式是负的, 则它有两个虚根.

证明 根据定理 7 的推论, 或者所有的根都是实根, 或者有两个共轭虚根 $z_1 = x_1 + y\mathrm{i}$ 和 $z_2 = x_1 - y\mathrm{i}$. 如果所有的根都是实的, 则对所有 $i \neq j$, 有 $(z_i - z_j)^2 \geqslant 0$, 因此 $D \geqslant 0$. 对第二种情况, 有 $(z_1 - z_2)^2 = -4y^2 < 0$, 又因为 $z_3 = x_3$ 是实的, 所以 $(z_1 - z_3)(z_2 - z_3) = (x_1 - x_3)^2 + y^2 > 0$, 因此 $D < 0$. 证毕

由 (23) 式, 条件 $D = 0$ 给出了检验方程有重根的简单判别法.

可惜的是, 在 $D > 0$ 的情况中, 方程 $z^3 + pz + q = 0$ 的三个根全部是实根, 但公式 (19) 却是用复数把它们表示出来. 我们在 15.6 节中将指出这是毫无助益的.

习 题

1. 证明: 对任意复数 y, p, 存在 z 满足 $y = z - \frac{p}{3z}$, 存在多少个 z?

2. 用根式表出方程的解：

(a)$z^2+\mathrm{i}z=2$，　(b)$z^3+3\mathrm{i}z=1+\mathrm{i}$，　(c)$z^3+3\mathrm{i}z^2=10\mathrm{i}$.

3. 把习题 2(a)~(c) 每个方程中的一个根改写成小数形式.

4. (a) 证明 (22) 式.　　　(b) 证明 (24) 式.

*5. (a) 证明：$\mathrm{sh}3\gamma=\mathrm{sh}(3\gamma+2\pi i)$.

(b) 利用 4.4 节中的公式 (9a) 证明：方程 $4z^3+3z=C$ 除有实根 $\mathrm{sh}\left[\frac{1}{3}\mathrm{arsh}C\right]=\mathrm{sh}\gamma$ 外, 还有复根 $-\frac{1}{3}\mathrm{ch}\gamma\pm\frac{\mathrm{i}\sqrt{3}}{2}\mathrm{sh}\gamma$.

6. 设 $\omega=c^{\frac{2\pi i}{5}}$ 是五次本原单位根, 且设 $\xi=\omega+\frac{1}{\omega}$.

(a) 证明：$\xi^2+\xi=1$.

(b) 推断：中心在 (0, 0), 一个顶点在 (1, 0) 的一个正五边形中, 与这个顶点相邻的顶点的 x 坐标是 $\frac{\sqrt{5}-1}{4}$.

7. 用公式 $\cos\theta=\frac{\mathrm{e}^{\mathrm{i}\theta}+\mathrm{e}^{-\mathrm{i}\theta}}{2}$ 证明：$\cos n\theta=T_n(\cos\theta)$, 其中 T_n 为一个适当的 n 次多项式, 并计算 T_1,T_2,T_3,T_4.

*5.6　四次方程的根式解法

任何一种把代数方程的求解化为一系列有理运算和对某数开 n 次方根的运算的方法称为“根式解法”.

定理 9　*任意 $n\leqslant 4$ 次实系数或复系数多项式方程可用根式求解.*

证明　因为 $n=1$ 的情形在任意域上都是可解的, 而 $n=2, 3$ 的情形在 5.5 节中已作了处理, 所以我们只须考虑

$$ax^4+bx^3+cx^2+dx+e=0,\quad (a\neq 0).$$

再有, 用 a 去除每一项, 并用 $z=x+\frac{b}{4a}$ 代替 x(以便配成“完全”四次方), 我们得到方程

$$z^4+pz^2+qz+r=0, \tag{25}$$

它的根与原方程的根相差 $\frac{b}{4a}$. 但是, 对所有的 u, (25) 式等价于

$$z^4+z^2u+\frac{u^2}{4}-z^2u-\frac{u^2}{4}+pz^2+qz+r=0, \tag{26}$$

$$\text{或}\left(z^2+\frac{u}{2}\right)^2-\left[(u-p)z^2-qz+\left(\frac{u^2}{4}-r\right)\right]=0.$$

第一项是一个完全平方 P^2, 这里 $P = z^2 + \frac{u}{2}$. 方括号中的项当选取 u 满足 (相当于判别式等于零)

$$q^2 = 4(u-p)\left(\frac{u^2}{4} - r\right) \tag{27}$$

时, 是一个完全平方 Q^2. 应用定理 8, 这个关于 u 的三次方程可用根式求解. 如果 (25) 式的系数是实数, 我们甚至可以证明, 至少有一个实数 $u_1 \geqslant p$ 满足 (27) 式, 这是因为, 当 $u = p$ 时, (27) 式的右边为零, 并且当 $u > 0$ 充分大时, (27) 式的右边大于 q^2, 或大于另一个任意预先给定的常数. 因此根据 4.4 节定理 4, (27) 式有所要求的实根 u_1.

把这个常数 u_1 代入 (26) 式, 则 (25) 式的左边采取形式 $P^2 - Q^2 = (P+Q)(P-Q)$, 或者

$$\left(z^2 + \frac{u_1}{2} + Q\right)\left(z^2 + \frac{u_1}{2} - Q\right), \tag{28}$$

这里

$$Q = Az - B, \qquad A = \sqrt{u_1 - p}, \qquad B = \frac{q}{2A}. \tag{29}$$

(25) 式的根显然是 (28) 式的两个二次因子的根, 后者可根据 (16) 式求出. 注意, 如果原方程的系数 a, b, c, d, e 都是实数, 那么这两个因子也是实系数的.

回顾一下方程根式解法的历史是有意义的. 二次方程的求解由印度人发现, 而它的几何形式由希腊人给出 (4.1 节). 三次方程和四次方程的求解是由文艺复兴时期意大利数学家 Scipio del Ferro (1515) 和 Ferrari (1545) 给出. 此外, 18 世纪末, 阿贝耳 (Abel) 和伽罗瓦 (Galois) 证明了所有次数 $n \geqslant 5$ 的多项式方程用根式求解是不可能的 (15.9 节).

习　题

1. 用根式求解 $z^4 - 4z^3 + (1+\mathrm{i})z = 3\mathrm{i}$.
2. 不用代数基本定理, 证明: 每个次数 $n < 6$ 的实系数多项式有复根.
3. 解联立方程

$$\begin{cases} zw = 1 + \mathrm{i}, \\ z^2 + w^2 = 3 - \mathrm{i}. \end{cases}$$

*5.7 稳定型方程

很多物理系统是稳定的当且仅当相应的多项式方程的全部根具有负的实部. 因此具有这种性质的方程称为 "稳定型" 方程.

在实二次方程 $z^2+Bz+C=0$ 的情形中, 容易检验它的稳定性. 如果 $4C\leqslant B^2$, 则两个根都是实数. 它们具有相同符号当且仅当 $z_1z_2=C>0$, 符号是负的当且仅当 $B=-(z_1+z_2)>0$. 如果 $4C>B^2$, 则方程的根是两个共轭复数, 它们两个具有负的实部 $x_1=x_2$ 当且仅当 $B=-2x_1=-2x_2>0$. 这种情形中也有 $C>\dfrac{B^2}{4}>0$. 因此这两种情形的 "稳定性" 条件是 $B>0,C>0$.

在实三次方程 $z^3+Az^2+Bz+C=0$ 的情形中, 稳定性条件也不难找到. (当然, 只考虑简化形式 (17) 还不够.) 事实上, 如果所有的根具有负实部, 那么, 因为一个根 $z=-a$ 是实的, 所以我们有分解式

$$z^3+Az^2+Bz+C=(z+a)(z^2+bz+c), \tag{30}$$

这里 $a>0$, 并由上述情况知 $b>0,c>0$. 因此稳定性的必要条件是 $A=a+b>0,B=(ab+c)>0$ 和 $C=ac>0$. 此外 $AB-C=b(a^2+ab+c)>0$.

反之, 假定 $A>0,B>0,C>0,AB-C>0$, 并考虑实分解式 (30), 根据定理 7 这个分解总是存在的. 因为 $ac=C>0$, 所以 a 和 c 有相同的符号. 但是, 如果它们两个都是负的, 那么 b 必须是负的才能使 $ab+c>0$, 因此 $A=a+b<0$, 同假定矛盾. 因此有 $a>0,c>0$, 并推出 $a^2+ab+c=a(a+b)+c>0$. 但是这就推出 $b=\dfrac{AB-C}{a^2+ab+c}>0$, 因此 (30) 式的两个因子是稳定的. 从而我们证明了下面结果.

定理 10 实二次方程 $z^2+Bz+C=0$ 是稳定型方程当且仅当 $B>0$ 和 $C>0$. 实三次方程 $z^3+Az^2+Bz+C=0$ 是稳定型方程当且仅当 $A>0,B>0,C>0$ 和 $AB>C$.

习 题

1. 检验下列多项式的稳定性:
   (a) z^3+z^2+2z+1, (b) z^3+z^2+2z+2.
2. 证明: n 次首一实系数多项式是稳定型的, 那么它的所有系数都必须是正的.

*3. 证明: 实系数多项式 $z^4+Az^3+Bz^2+Cz+D$ 是稳定型的当且仅当它的所有系数都是正的, 并且 $ABC>A^2D+C^2$.

*4. 利用习题 3, 求出复系数二次方程 $z^2+Bz+C=0$ 是稳定型方程的充分必要条件.(提示: 考虑 $(z^2+Bz+C)(z^2+B^*z+C^*)=0$.)

第6章　群

6.1　正方形的对称

"对称"的概念对每个受过教育的人来说都是熟悉的，但是由对称产生的对称代数却只有少数人了解. 我们将通过具体的正方形对称来引出这个代数.

我们设想一个正方形硬纸板放在有固定轴的平面上，使得正方形的中心落在坐标原点上，正方形的一个边是水平的. 显然，这个正方形具有旋转对称：它通过下面的刚体运动可旋转成自身.

R：围绕中心 O 顺时针旋转 $90°$.

R', R''：以同样的方式旋转 $180°$ 和 $270°$.

这个正方形还有反射对称：它可以通过下面的刚体反射变为自身.

H：关于过原点 O 的水平轴的反射.

V：关于过原点 O 的垂直轴的反射.

D：关于Ⅰ，Ⅲ象限中的对角线的反射.

D'：关于Ⅱ，Ⅳ象限中的对角线的反射.

至此，我们列举的这些情形包括了七种对称.

对称代数起源于下述事实；我们通过相继完成两个运动可以把两个运动相乘. 例如，乘积 HR 可分两步得到：首先把正方形关于水平轴反射，然后再把正方形顺时针旋转 $90°$. 通过正方形硬纸板的实验，我们可以验证，HR 的最终效果与 D' 是一样的，这里 D' 是关于从左上角到右下角的对角线的反射. 另一方面，等式 $HR = D'$ 可以通过观察正方形的每个顶点的变化来验证，如果等式两边具有同一个效果，则等式成立. 例如，在图 1 中，HR 是先通过 H 把 1 送到 4，然后通过 R 把 4 送到 3，因此就把 1 送到 3，这恰好与 D' 的效果一样.

类似地，RH 定义为先顺时针旋转 $90°$ 随后关于水平轴反射. (注意：图 6-1 的平面包含反射轴，这个平面可以想象成不随正方形而旋转.)

由计算表明 $RH = D \neq HR$，由此我们顺便得到这里所说的"乘法"一般不满足交换律！但是它满足结合律，我们在 6.2 节中将看到这一点.

读者计算正方形对称的其他乘积 (6.4 节的表 1 中给出一个完整的乘法表) 是有意义的. 当你做完这些乘积之后将会发现，一般地，逐次地把任意两个对称乘起来便得到第三个对称，但有个例外，例如，当 R 和 R'' 相乘时，就会看到它的积是一个使正方形每个点都保持原来位置的运动，这就是所谓的"恒等"运动 I. 这通常不被

非数学家认为是对称; 尽管如此, 为了能使所有的对称两两相乘, 我们还是把 I 看作一个 (退化的) 对称.

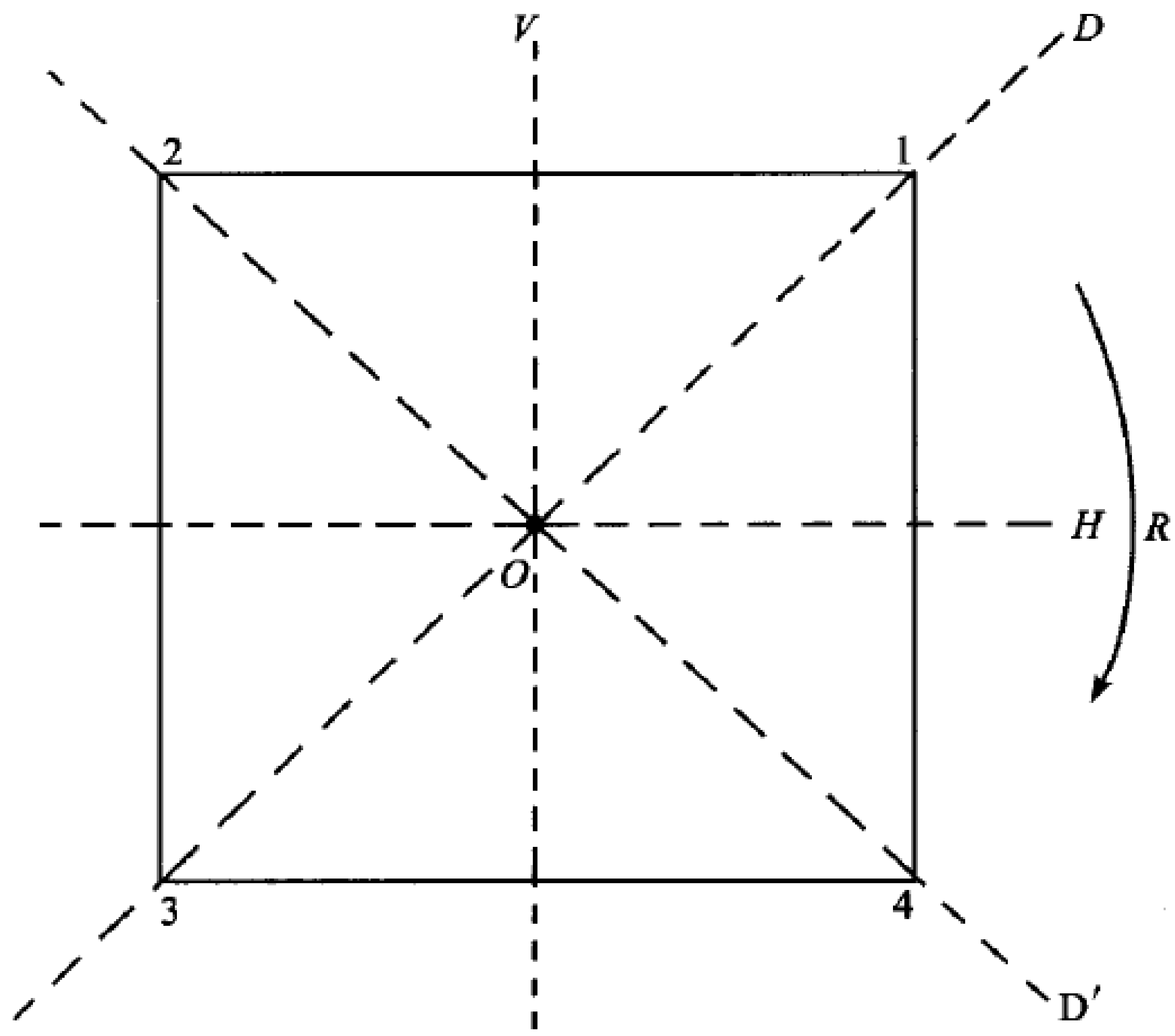


图 6-1

一般地, 根据定义, 几何图形的对称是图形上的点保持距离不变的一一变换. 容易看出, 正方形的任意对称一定把顶点 1 变换到四个可能的顶点之一, 而且对每个这样的选择正好有两个对称. 于是总共只有八个对称, 就是我们已经列出来的那些.

不仅正方形, 而且每个正多边形和正多面体 (例如立方体和正二十面体) 都存在有趣的对称群, 可以用上面概述的初等方法找到.

类似地, 很多装饰品有有趣的对称. 例如我们考虑一个无限长的装饰图案



在这个图案中, 箭头是沿着直线以一英寸间隔均匀分布的. 这个图形的三个简单的对称是: T, 向右平移一英寸; T', 向左平移一英寸; H, 图形关于水平轴反射. 其他对称 (事实上是一切对称) 可以由这三个对称反复相乘而得到.

习 题

1. 计算 $HV, HD', D'H, R'D', D'R', R'R''$.
2. 在 "箭头" 装饰图案中, 描述对称 TH 和 HT.
3. 列出等边三角形的所有对称, 并计算五种具有代表性的乘积.
4. 列出普通矩形的所有对称, 并计算它们的所有乘积.

*5. 正四面体有多少对称? 正八面体有多少对称? 画图说明.

*6. 证明: 正文中的装饰图案的任意对称可通过 H, T 和 T' 反复相乘而得到.

6.2 变 换 群

对称代数可以推广到无论什么元素的任意集合 S 的一一变换. 虽然常常把集合 S 看作"空间"(例如平面或球), 把 S 的元素看作"点", 并且把双射看作 S 的相对适当性质的"对称", 然而在任何情况下, S 的双射还满足一些非平凡的代数定律.

为了理解这些定律, 我们必须清楚地记住 1.11 节中给出的关于函数、单射、满射和双射的定义. 为了重新解释它们, 我们给出一些新的例子. 同 1.11 节中一样, 我们通常用缩写记号 xf 代替 $f(x)$(读作"x 通过 f 的变换"), 用 xg 代替 $g(x)$, 等等.

函数 $f(x) = \mathrm{e}^{2\pi i x}$ 把实数域 $\mathbf{R}$ 映入复数域 $\mathbf{C}$, 它的值域 (像) 是单位圆. 类似地, $g(z) = |z|$ 是函数 $g: \mathbf{C} \to \mathbf{R}$, 它的像是所有非负实数的集合.

再有, 考虑下列整数环 $\mathbf{Z}$ 到自身的函数 $\phi_0 : \mathbf{Z} \to \mathbf{Z}$ 和 $\psi_0 : \mathbf{Z} \to \mathbf{Z}$:

$$n\phi_0 = 2n, \quad m\psi_0 = \begin{cases} \dfrac{m}{2}, & \text{当 } m \text{ 为偶数}, \\ 0, & \text{当 } m \text{ 为奇数}. \end{cases}$$

根据乘法消去律, ϕ_0 是一一的; 然而它的值域仅由偶数组成, 所以 ϕ_0 没有把 $\mathbf{Z}$ 变换到 $\mathbf{Z}$ 上. 另一方面, ψ_0 不是一一的, 这因为所有奇数都映射到零, 但是它把 $\mathbf{Z}$ 映到 $\mathbf{Z}$ 上, 于是 ψ_0 是满射, 而不是单射.

我们现在转到变换代数. 具有相同的定义域 S 和相同的取值域 T 的两个变换 $\phi : S \to T$ 和 $\phi' : S \to T$, 如果它们作用到 S 的每一点上都有相同的效果, 则称它们**相等**, 即

$$\phi = \phi' \text{的意思是, 对每个} p \in S, \text{有} p\phi = p\phi'. \tag{1}$$

再定义两个变换 ϕ 和 ψ 的**乘积或合成** $\phi\psi$ 为它们相继作用的结果, 先 ϕ 后 ψ, 然而这里应假定 ϕ 的取值域是 ψ 的定义域. 换句话说, 如果

$$\phi : S \to T, \qquad \psi : T \to U,$$

那么 $\phi\psi$ 是由等式

$$p(\phi\psi) = (p\phi)\psi \tag{2}$$

给出的由 S 到 U 中的变换, 式中规定 $\phi\psi$ 作用到任意点 $p \in S$. 特别是, S(到自身) 的两个变换的乘积总是可以定义的. 我们现在只考虑这种情况, 只要假定所包含的乘积有定义, 下面证明的恒等式几乎所有都可以用于一般情况.

变换的乘法适合

结合律　　$(\phi\psi)\theta = \phi(\psi\theta)$,

这里假定所包含的乘积都有定义. 直观上这是显然的: $(\phi\psi)\theta$ 和 $\phi(\psi\theta)$ 两者都是按照先 ϕ 后 ψ 最后 θ 的顺序作用的. 正式地, 对每个 $p \in S$, 我们有

$$p[\phi(\psi\theta)] \underset{\phi(\psi\theta)}{=} (p\phi)(\psi\theta) \underset{\psi\theta}{=} [(p\phi)\psi]\theta \underset{\phi\psi}{=} [p(\phi\psi)]\theta \underset{(\phi\psi)\theta}{=} p[(\phi\psi)\theta],$$

这里每步都依赖于乘法的定义 (2), 也就是把定义 (2) 用到与每步相对应的等号下面标出的乘积上. 根据变换相等的定义 (1), 这就证明了结合律 $\phi(\psi\theta) = (\phi\psi)\theta$.

集合 S 上的*恒等变换* $I = I_s$ 是使 S 上每个点保持固定的变换 $I : S \to S$. 代数上, 这可叙述成等式

$$pI = p, \qquad 对每个 p \in S. \tag{3}$$

从上面的定义, 直接推出

同一律　　$I\phi = \phi I = \phi$,　　对一切ϕ.

为了验证这一点, 我们注意, 对所有的 p, 有 $p(I\phi) = (pI)\phi = p\phi$, 类似地, $p(\phi I) = (p\phi)I = p\phi$.

现在回到前面定义在集合 $\mathbf{Z}$ 上的特殊变换 ϕ_0 和 ψ_0, 并计算它们的乘积. 显然

$$m\psi_0\phi_0 = \begin{cases} m, & 当m为偶数, \\ 0, & 当m为奇数. \end{cases}$$

因此 $\psi_0\phi_0 \neq I$. 另一方面, 对一切 $m \in \mathbf{Z}$, 有 $m\phi_0\psi_0 = m$, 因此 $\phi_0\psi_0 = I$. 于是我们称 ψ_0 是 ϕ_0 的右逆元素 (而不是左逆元素).

一般地, 如果变换 $\phi : S \to S$ 和 $\psi : S \to S$ 具有 $\phi\psi = I : S \to S$, 那么称 ϕ 是 ψ 的*左逆元素*, 而 ψ 是 ϕ 的*右逆元素*. 这些定义同以前定义的是 "一一的 (单射)" 和 "映上的 (满射)" 等概念有密切关系.

定理 1　*变换 $\phi : S \to S$ 是一一的当且仅当它有右逆元素, ϕ 是映上的当且仅当它有左逆元素.*

证明　如果 ϕ 有右逆元素 $\psi, \phi\psi = I$, 并且 $p\phi = p'\phi$, 那么

$$p = p(\phi\psi) = (p\phi)\psi = (p'\phi)\psi = p'(\phi\psi) = p'.$$

于是由 $p\phi = p'\phi$ 可推出 $p = p'$, 因此 ϕ 是一一的. 类似地, 如果 ϕ 有左逆元素 ψ', 则 $\psi'\phi = I$. 因此 S 中的任何元素 q 都可写成

$$q = qI = q(\psi'\phi) = (q\psi')\phi,$$

这表明 q 是某一点 $p=q\psi'$ 的 ϕ- 像. 因此 ϕ 是映上的.

反过来, 已知任意$\phi: S\to S$, 我们首先如下构造第二个变换 $\psi: S\to S$. S 中有一些点, 其中每个点 q 是 S 的一个或多个点 p 在 ϕ 之下的像, 对每个点 q, 在这些点 p 中任意选出[①]一个点作为像 $q\psi$. 那么, 对形为 $p\phi$ 的任何一个 q, 有

$$q(\psi\phi)=(q\psi)\phi=p\phi=q.$$

再令 ψ 随便按什么方式映射 S 中其余的点 q, 譬如说映射到 (非空) 集合 S 的某个固定点上.

现在, 如果 ϕ 是映上的, 那么每个 q 都有形式 $p\phi$, 因此 $\psi\phi=I$, 所以 ϕ 有 ψ 作为它的左逆元素. 另一方面, 如果 ϕ 是一一的, 那么, 对每个 $p,(p\phi)\psi$ 一定是唯一的 p, 即上面所说的 $q=p\phi$ 中的 p. 因此 $\phi\psi=I$, 所以 ψ 是 ϕ 的右逆元素, 如断言所述.

证毕

注 微积分学中函数记号 $y=\phi(x)$ 暗示记成 $y=\phi x$, 而前面我们写成 $y=x\phi$; 按照这种记号, ϕ 和 $z=\psi(y)$ 的合成自然写成 $z=(\psi\phi)x$, 它是作为 $z=\psi(\phi(x))$ 的缩写记号, 并代替 $z=x\phi\psi$. 因此 $\psi\phi$ 的意思是"先执行 ϕ, 后执行 ψ", 而右逆元素和左逆元素的概念应相互对换. 虽然上述两种记号用任何一种都是可以的, 但是一定要避免它们之间的混淆. 然而, 双边逆元素的意思保持不变, 正如下面推论所述.

推论 1 *变换 $\phi: S\to S$ 是双射当且仅当它既有右逆元素又有左逆元素. 如果 ϕ 是双射, 那么 ϕ 的任意右逆元素等于 ϕ 的任意左逆元素.*

事实上, 如果 ϕ 有右逆元素 θ 和左逆元素 ψ, 那么

$$\theta=I\theta=(\psi\phi)\theta=\psi(\phi\theta)=\psi I=\psi.$$

把变换 $\phi: S\to S$ 的 (双边) 逆元素定义为满足

逆律 $$\phi\phi^{-1}=\phi^{-1}\phi=I$$

的任意变换 ϕ^{-1}. 这些等式也表明 ϕ^{-1} 是 ϕ 的 (双边) 逆元素, 因此进一步有

推论 2 *变换 $\phi: S\to S$ 是双射当且仅当 ϕ 有 (双边) 逆元素 ϕ^{-1}. 如果 ϕ 是双射, 那么 ϕ 的任何两个逆元素是相等的, 并有*

$$(\phi^{-1})^{-1}=\phi. \tag{4}$$

这个推论后面将要用到. 它可以直接证明, 因为 ϕ^{-1} 只不过是这样一个变换, 它把 S 的每个点 $q=p\phi$ 变回原来唯一的点 p. 在 S 是有限的特殊情况下, ϕ 是一一的当且仅当 ϕ 是映上的, 因此在这种情况下左逆元素和右逆元素的更细致的讨论是没有意义的.

① 在这样的点 q 组成的集合是无限的情况下, 选择公理(参见 12.2 节) 断言: 对每个 q, 可以选择无限多个这样的 p.

对于集合 S 到另一个集合 T 的函数 $\phi: S \to T$ 来说, 定理 1 及其推论以及它们的证明也都成立. 我们只须注意, 左逆元素 ψ 或者右逆元素 θ 是第二个集合 T 到集合 S 中的变换, 并注意.

$$\psi\phi = I_T : T \to T, \qquad \phi\theta = I_S : S \to S.$$

这里 I_S 和 I_T 分别是 S 和 T 上的恒等变换.

我们现在准备定义变换群这一重要概念. "空间" S 上的*变换群*是指满足下列条件的把 S 映上 S 的一一变换 ϕ 组成的任意集合 G:

(i) S 的恒等变换在 G 中;

(ii) 如果 ϕ 在 G 中, 则它的逆元素也在 G 中;

(iii) 如果 ϕ 和 ψ 在 G 中, 则它们的积也在 G 中.

定理 2 *任意空间 S 到自身的所有双射所组成的集合 G 是一个变换群.*

证明 因为 $II = I, S$ 上的恒等变换 I 是双射, 因此 I 在集合 G 中, 上面的条件 (i) 满足. 如果 ϕ 在 G 中, 由前面的推论 2 得 ϕ^{-1} 也是双射, 因此它同样在 G 中, 条件 (ii) 满足. 最后, 任意两个把 S 映上 S 的一一变换 ϕ 和 ψ, 它们的乘积有逆元素, 因为根据假设

$$(\phi\psi)(\psi^{-1}\phi^{-1}) = \phi(\psi\psi^{-1})\phi^{-1} = \phi I\phi^{-1} = \phi\phi^{-1} = I,$$

$$(\psi^{-1}\phi^{-1})(\phi\psi) = \psi^{-1}(\phi^{-1}\phi)\psi = \psi^{-1}I\psi = \psi^{-1}\psi = I.$$

因此 $\phi\psi$ 也是双射, 并且有逆元素

$$(\phi\psi)^{-1} = \psi^{-1}\phi^{-1}. \tag{5}$$

口头上说就是, 乘积的逆元素等于逆元素颠倒次序相乘. 证毕

有限集 S 到它自身的双射通常称为 S 的*置换*. n 个元素的所有置换组成的群称为 *n 次对称群*; 显然它包含 $n!$ 个置换, 这因为第一个元素的像 k_1 可以有 n 种方式选取, 然后, 第二个元素的像可以从去掉 k_1 剩下的元素中以 $n-1$ 种方法选取, 等等.

习 题

1. 在正方形对称群中计算 $VD, (VD)R'', DR'', V(DR'')$.
2. 类似习题 1, 计算 $HR, R'(HR), R'H, (R'H)R$.
3. 设 S 由所有实数组成 (或由直线上的所有点 x 组成), 所考虑的变换具有形式 $x\phi = ax + b$. 在下列各种情况中, 以所指定类型的 a 和 b 为系数的所有可能的变换 ϕ 组成的集合, 哪些是变换群, 并给出证明.

(a) a 和 b 是有理数; (b) $a=1, b$ 是奇数;
(c) $a=1, b$ 是正整数或零; (d) $a=1, b$ 是偶数;
(e) a 是整数, $b=0$; (f) $a\neq 0, a$ 和 b 是实数;
(g) $a\neq 0, a$ 是整数, b 是实数; (h) $a\neq 0, a$ 是实数, b 是整数;
(i) $a\neq 0, a$ 是整数, b 是无理数; (j) $a\neq 0, a$ 是有理数, b 是实数.

在这些变换群中, 哪些群的乘法满足交换律?

4. 找出恰有三个"点"的"空间" S 上的所有变换, 共有多少个变换? 其中有多少个是一一变换?
5. 证明: 所有正整数的集合上的变换 $n\longmapsto n^2$ 没有左逆元素. 并列出两个明显的右逆元素.
6. 列出正文中定义的变换 $\psi_0:\mathbf{Z}\to\mathbf{Z}$ 的两个不同的左逆元素, 并列出 ϕ_0 的两个不同的右逆元素.
7. 证明: 如果 ϕ 和 ψ 都有右逆元素, 那么 $\phi\psi$ 也有右逆元素.
8. 对于正方形对称群, 计算 $[R^{-1}(VR)]^{-1}[(R^{-1}D)R]$.
9. 对正方形对称群, 解方程 $RXR'=D$.
10. 在正方形对称群中, 验证

$$(RH)^{-1}=H^{-1}R^{-1}\neq R^{-1}H^{-1}.$$

11. 求出矩形的每个对称的逆元素, 并验证公式 (5).
12. 证明: 如果 $\phi_1,\phi_2,\cdots,\phi_n$ 是一一的, 那么 $\phi_1\phi_2\cdots\phi_n$ 也是一一的, 且有逆元素

$$(\phi_1\phi_2\cdots\phi_n)^{-1}=\phi_n^{-1}\cdots\phi_2^{-1}\phi_1^{-1}.$$

13. 证明: 对任意 $\phi:S\to S$, 由定理 1 证明的第二部分所构造的变换 ψ 满足 $\phi\psi\phi=\phi$.

*14. 证明: 具有唯一右逆元素或唯一左逆元素的变换 $\phi:S\to S$, 必是 S 到 S 上的一一变换.

6.3 其他例子

立方体的所有对称构成另一个有趣的群. 用几何语言来说, 这些对称是保持立方体上距离不变的一一变换. 它们被称为"等距变换", 共有 48 个. 为了说明这一点, 我们注意到, 任意一个初始顶点可以变换到八个顶点中任意一点, 任意顶点的变换固定之后, 这个顶点的三个相邻顶点可以有六种方式进行排列, 于是给出 6·8=48 种可能性. 当一个顶点和它的三个相邻顶点的位置确定时, 立方体上任何一点的位置也就固定下来, 所以整个对称就知道了. 因此立方体恰有 48 个对称. 它们中间很多都具有特殊的几何性质, 例如, 其中一个对称是把立方体的每个点反射成对径点.

包含着无穷多个变换的一个熟悉的群是所谓欧几里得群. 这个群由平面的所有 "等距变换" 组成, 或者用初等几何的语言来说, 在这些变换下, 平面同自身是全等的. 这个群由平移、刚体旋转和反射的乘积组成. 我们将在第 9 章详细讨论它.

另一个群是由空间的所有 "相似变换" 组成, 即由那些使一切距离扩大常数 $k(k > 0$, 称为比例因子) 倍的一一变换组成. 任意球面变为自身的所有刚体运动又构成一个群. 使平面上正六边形网络 (图 6-2) 保持不变的所有 "等距变换" 构成另一个有趣的群.

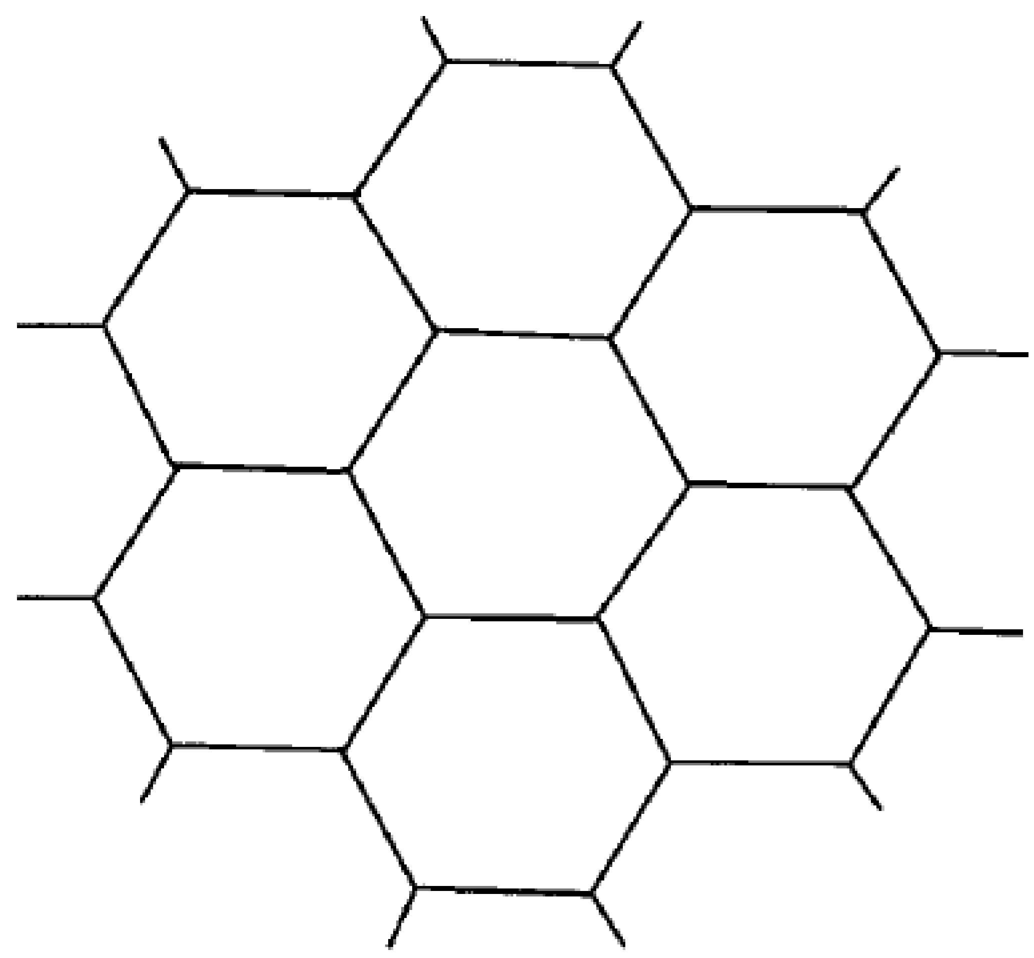


图　6-2

再有, 一条橡皮绳沿一直线摆放着, 绳的两端分别固定在 P, Q 两点, 它可以沿着这条直线以很多种方式变形. 所有这些变形构成一个群 (通常称为线段 PQ 的同胚群).

一般地说, 任意集合的一一变换, 如果保持集合中元素的某个或某些任意给定的性质, 那么这些一一变换构成一个群. 克莱茵 (Felix Klein)(Erlanger Programm, 1872) 雄辩地描述了, 不同的几何分支可以看作是研究相应空间的那些在适当的变换群下保持不变的性质. 例如, 欧几里得几何是研究空间的那些在所有等距变换下保持不变的性质, 拓扑学是研究空间的那些在所有同胚变换之下保持不变的性质. 类似地, 射影几何和仿射几何分别研究空间在射影群和仿射群下保持不变的性质. 射影群和仿射群的定义将在第 9 章给出.

习　题

1. 描述带有六个等间隔辐条的车轮的全部对称.
2. 描述一个顶点固定的立方体的六个对称.
3. 设 S,T 是立方体关于两个平面的反射, 这两个平面分别平行于立方体的两个不同的侧

面. 描述 ST 的几何意义.

4. 描述一些把图 2 的正六边形网络变到自身的平面等距变换.
5. 对正方形网络做习题 4. 你能数出所有这样的变换吗 (这是困难的)?
6. 对正三角网络做习题 4, 并说明这些变换与习题 1 的变换群的关系.
7. 对下述几种情况做习题 4:

   (a) 无限圆柱体, (b) 有限圆柱体,

   (c) 圆柱螺旋线, 即一条围线柱面并与圆柱轴线成定角的螺旋线.

*8. 证明: 所有变换 $x \longmapsto x' = \dfrac{ax+b}{cx+d}$(其中系数 a,b,c,d 在任意域 F 中, 并且 $ad-bc=1$) 组成一个群, 这些变换作用在由域 F 的全体元素和符号元素 ∞ 组成的集合上.

6.4 抽 象 群

变换群决不是其乘法满足 6.2 节中所说的结合律、同一律和逆律的唯一系统. 例如, 任意域 (如有理数域、实数域和复数域) 的全体非零元素都满足这些定律. 任意两个非零元素的乘积是一个非零元素; 结合律成立; 域的单位元素 1 满足同一律, 并且 $\dfrac{1}{x}=x^{-1}$ 满足逆律.

类似地, 任意整环的全体元素 (这次包括零) 在加法运算之下满足上述三个定律. 例如, 任意两个元素有唯一确定的和; 加法满足结合律; 对于加法运算, 零满足同一律, $-x$ 满足逆律. 换句话说, 任意整环的全体元素在加法之下构成一个群.

为方便起见, 我们引进包含上述和其他一些例子的群的抽象概念.

定义 具有二元运算的元素集合 G,(i) 运算满足结合律; (ii) 有一个满足同一律的单位元素; (iii) 对每个元素 a, 有元素 a^{-1}(称为 a 的逆) 满足逆律, 则这个集合 G 称为群.

我们可以不提变换, 用许多方式抽象地给出群的定义, 这样定义的群常常称为抽象群.

在讨论抽象群的时候, 元素用小写拉丁字母 $a,b,c,\cdots$ 来表示. 乘积记号 "ab" 通常用来表示 G 的两个元素 a 和 b 在群的运算之下而得的结果, 但是其他记号, 像 "$a+b$" 和 "$a\circ b$" 也同样适用. 在乘积记号中, 用 "e" 表示单位元素, 定义群的三个定律变为

结合律 $a(bc)=(ab)c$, 对一切 a,b,c.

同一律 $ae=ea=a$, 对一切 a.

逆 律 $aa^{-1}=a^{-1}a=e$, 对每个 a 和某个 a^{-1}.

其运算满足交换律的群称为交换群或阿贝耳群, 利用这个概念我们可以把域的定义简化如下.

定义　集合 F 满足下列条件时称为域, F 在两个唯一确定的二元运算 —— 加法和乘法之下是封闭的, 并有

(i) 在加法之下, F 是具有单位元素零的交换群;

(ii) 在乘法之下, F 中非零元素构成交换群;

(iii) 分配律成立：$a(b+c)=ab+ac$.

为证明这个定义同 2.1 节中给出的定义是等价的, 我们注意, 这里给出的公设, 除了含有因子零的乘法结合律外, 包含前面对域所描述的一切公设. 这可以详细地验证.

1.1 节和 1.2 节中的一些结果现在将表现为下面关于群的定理的推论.

定理 3　在任意群中, $xa=b$ 和 $ay=b$ 有唯一解, 分别为 $x=ba^{-1}$ 和 $y=a^{-1}b$. 因此由 $ca=da$ 可推出 $c=d$, 同样由 $ac=ad$ 可推出 $c=d$(消去律).

证明　如果 a^{-1} 是在逆律中确定的元素, 显然, $(ba^{-1})a=b(a^{-1}a)=be=b$, 类似地, $a(a^{-1}b)=b$. 反过来, 由 $xa=b$ 可推出 $x=xe=xaa^{-1}=ba^{-1}$, 类似地, 由 $ay=b$ 可推出 $y=a^{-1}b$.

注意, 在这个证明中并没有假定 a^{-1} 是满足 $xa=e$ 的唯一的元素. 但 a^{-1} 确是唯一的, 这是因为, 若 $xa=e$, 则

$$x=xe=x(aa^{-1})=(xa)a^{-1}=ea^{-1}=a^{-1}.$$

类似地, a^{-1} 是使得 $ay=e$ 的唯一元素.

因为根据定理 3, 在任意群 G 中, 方程 $ex=e$ 和 $ay=e$ 有唯一解, 分别为 $x=e$ 和 $y=a^{-1}$, 因此我们得到

推论　群有唯一的单位元素, 并且对每个元素 a 有唯一的逆 a^{-1}.

定理 4　前面所述的群的定义中, 同一律和逆律可以用较弱的形式来代替.

左同一律　对所有的元素 a, 存在某元素 e, 满足 $ea=a$.

左 逆 律　对给定的元素 a, 存在某元素 a^{-1}, 满足 $a^{-1}a=e$.

证明　如果这些弱的定律成立, 则左消去律也成立, 即由 $ca=cb$ 可推出 $a=b$, 因为我们只须用 c^{-1} 左乘等式 $ca=cb$ 的两边, 再用结合律得到 $(c^{-1}c)a=(c^{-1}c)b$, 这就是 $ea=eb$, 故得 $a=b$.

给出的这个左单位元素也是右单位元素, 这是因为

$$a^{-1}ae=ee=e=a^{-1}a,$$

再根据左消去律, 因此对所有的 a, 有 $ae=a$. 最后, 左逆元素也是右逆元素, 因为由于左单位元素也是右单位元素, 则有

$$a^{-1}(aa^{-1})=(a^{-1}a)a^{-1}=ea^{-1}=a^{-1}=a^{-1}e,$$

现在再用左消去律, 得 $aa^{-1}=e$. 这就完成了我们的证明. 证毕

还有很多其他的群公设系统, 常用的一个是按照除法的可能性来建立的, 如下所述:

定理 5 如果 G 是一个非空集合, 在满足结合律的乘法之下是封闭的, 对于这个集合所有的方程 $xa=b$ 和 $ay=b$ 在 G 中有解 x 和 y, 那么 G 是一个群

证明留作习题 (习题 12).

除了对任意群 G 把有关乘法的代数定律系统化以外, 当 G 的元素有限时, 我们还可以用 "乘法表" 的形式给出 G 中任意两个元素乘积的特殊构成法则. 这个乘法表是一些元素的正方形阵列, 表的最左一列和最上一行列出群的所有元素. 表中对应着最左列上的 a 和最上行的 b 的那个元素是乘积 ab(按此次序).

为举例说明, 我们在表 1 中绘制了正方形对称群的乘法表. 这个表的计算可以按照 6.1 节中证明的 $HR=D'$ 和 $RH=D$ 的模式来进行. 其他方法将在 6.6 节中描述.

表 1 正方形对称群

	I	R	R'	R''	H	V	D	D'
I	I	R	R'	R''	H	V	D	D'
R	R	R'	R''	I	D	D'	V	H
R'	R'	R''	I	R	V	H	D'	D
R''	R''	I	R	R'	D'	D	H	V
H	H	D'	V	D	I	R'	R''	R
V	V	D	H	D'	R'	I	R	R''
D	D	H	D'	V	R	R''	I	R'
D'	D'	V	D	H	R''	R	R'	I

关于群的大部分性质可以直接从表中看到. 例如, 单位元素的存在表明, 某一行和相应的列一定分别是顶头一行和最左边一列的复制品. 方程 $ay=b$ 可解意味着 a 所在的那一行一定包含元素 b; 因为解是唯一的, 所以 b 在这一行中只能出现一次. 一个群是交换群当且仅当它的乘法表关于主对角线 (即左上角到右下角的联线) 是对称的. 遗憾的是, 结合律不容易从这个表中直观地看出.

习 题

1. 设 a,b,c 是群的固定元素, 证明方程 $xaxba=xbc$ 有唯一解.
2. 证明: 在 $2n$ 个元素的群中, 除单位元素外还存在一个元素同它的逆相等.
3. 全体正实数在加法下构成一个群吗? 在乘法下构成群吗? 全体偶数在加法下构成群吗? 全体奇数呢? 为什么?
4. 在模 11 整数域 $\mathbf{Z}_{11}$ 中, 下列集合中哪些在乘法下构成群:
   (a) (1, 3, 4, 5, 9), (b) (1, 3, 5, 7, 8), (c) (1, 8), (d) (1, 10).

5. 证明：含有四个元素或少于四个元素的群一定是阿贝耳群. (提示：ba 是 e, a, b, ab 中的一个, 平凡的情形除外.)

6. 证明：如果在一个群中 $xx = x$, 则 $x = e$.

7. 下列乘法表描述一个群吗？

	a	b	c	d
a	b	d	a	c
b	d	c	b	a
c	a	b	c	d
d	c	a	d	b

	a	b	c	d
a	a	b	c	d
b	b	a	d	c
c	c	d	a	a
d	d	c	b	b

8. 证明：1.2 节中法则 2、法则 4 和法则 6 在任意交换群中都成立.

9. 下列数集中哪一些是群？为什么？

(a) 所有有理数, 在加法运算之下; 在乘法运算之下.

(b) 所有无理数, 在乘法运算之下.

(c) 所有绝对值为 1 的复数, 在乘法运算之下.

(d) 所有绝对值为 1 的复数, 在运算 $z \circ z' = |z| \cdot z'$ 之下.

(e) 所有整数, 在减法运算之下.

(f) 任意整环的全体单位 (3.6 节), 在乘法运算之下.

10. 证明：下列公设系统描述一个阿贝耳群：

(i) 对一切 a, b, c 有 $(ab)c = a(cb)$;

(ii) 定理 4 的"左同一律"成立;

(iii) 定理 4 的"左逆律"成立.

*11. 证明：如果对群 G 中所有元素都有 $x^2 = e$, 那么 G 是交换群.

*12. 证明定理 5. (提示：如果 $ax = a$, 那么 x 是右单位元素, 并且任意右单位元素等于左单位元素.)

*13. 设 S 是一个非空集合, 在乘法运算之下是封闭的, 并且满足 $ab = ba$, $a(bc) = (ab)c$, 由 $ax = ay$ 可推出 $x = y$.

(a) 证明：若 S 有限, 则 S 是群.

(b) 证明：若 S 有限或无限, 则 S 可以嵌入一个群中.

6.5 同 构

考虑实数整环上的变换 $x \longmapsto \ln x$. 我们知道, 当 x 在区间 $0 < x < +\infty$ 上增加时, $\ln x$ 就在区间 $-\infty < x < +\infty$ 上连续增加; 也就是说, 这个对应是正实数系和全体实数系之间的一一对应 (逆变换是 $y \longmapsto e^y$). 而且对所有的 x, y, 有 $\ln(xy) = \ln x + \ln y$, 于是我们可以用相应的和的计算代替乘积的计算. 事实上, 这是对数主要的实际用途.

其次, 设 $\mathbf{Z}_3$ 是模 3 整数构成的域 (3.10 节), 并设 G 是等边三角形到自身的刚体旋转群. 如果 I, R 和 R' 分别表示转过 $0°, 120°$ 和 $240°$ 的旋转, 那么把整数同旋转联系起来的对应 $0 \leftrightarrow I, 1 \leftrightarrow R, 2 \leftrightarrow R'$ 是一个把 $\mathbf{Z}_3$ 中元素的和映射成 G 中相应旋转的乘积的双射. 例如, 考虑对应

$$\begin{aligned} 1+2 \equiv 0 (\mathrm{mod}\ 3) \quad &\leftrightarrow \quad RR' = I, \\ 2+2 \equiv 1 (\mathrm{mod}\ 3) \quad &\leftrightarrow \quad R'R' = R. \end{aligned}$$

这些都是 1.12 节中所谈到的"同构"一般概念的例子. 这个概念对群来说比对整环更简单也更重要.

定义 *两个群 G 和 G' 之间的同构指的是它们元素之间保持群的乘法的双射 $a \leftrightarrow a'$, 即它满足, 若 $a \leftrightarrow a'$ 和 $b \leftrightarrow b'$, 则 $ab \leftrightarrow a'b'$.*

例如, 在第一个例子中我们描述了正实数乘法群与实数加法群之间的同构. 在第二个例子中, 我们指出一个模 3 整数加法群与正三角形旋转对称群之间的同构.

类似地, 映射 $0 \longmapsto 1, 1 \longmapsto 2, 2 \longmapsto 4, 3 \longmapsto 3$ 是模 4 整数加法群与模 5 非零整数乘法群之间的同构. 通过比较模 4 整数加法群的加法表和模 5 非零整数乘法群的乘法表来验证这个结果是方便的. 见表 2 和表 3.

依次我们有, 模 4 整数加法群同构于正方形旋转对称群. 通过比较表 2 和表 1(6.4 节) 的旋转部分可以验证, 双射 $0 \leftrightarrow I, 1 \leftrightarrow R, 2 \leftrightarrow R', 3 \leftrightarrow R''$ 是一个同构.

表 2

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

表 3

×	1	2	4	3
1	1	2	4	3
2	2	4	3	1
4	4	3	1	2
3	3	1	2	4

同构的概念很重要, 因为它使我们认识到, 完全不同内容的群从抽象群论的观点看可以看成同一个群. 同构的群抽象地认为是同一个群 (它们的差别仅在于它们元素符号的不同), 这个事实可以在很多情况下看到.

例如, 根据定义, 两个有限群 G 和 G' 同构当且仅当通过适当的替换, 从 G 的群表可以得出 G' 的群表. 从 6.4 节的倒数第二句可以得出, G' 是阿贝耳群当且仅当 G 是阿贝耳群, 也就是说, 一个有限阿贝耳群的任何同构像是阿贝耳群. 还有, 在其他方面, 同构的性质很像相等.

定理 6 *关系"群 G 同构于群 G'"满足群之间的自反的、对称的和传递的关系.*

证明 自反性是显然的 (每个群通过恒等变换同它自身同构). 对于对称性, 设 $a \leftrightarrow aT$ 是 G 和 G' 之间的任意同构对应, 因为 T 是双射, 所以 T 有逆元素 T^{-1}, T^{-1} 是

G' 到 G 上的同构. 最后, 如果 T 把 G 同构地映射到 G' 上, 而 T' 把 G' 同构地映射到 G'' 上, 那么 TT' 就是 G 和 G'' 之间的同构. 证毕

值得注意的是, 定理 6 及其证明对于整环之间的同构同样成立, 而且对于任何类型的代数系统之间的同构也都成立.

定理 7 *在两个群同构之下, 它们的单位元素相互对应, 相应元素的逆元素相互对应.*

证明 方程 $ax = a$ 的唯一解 e 对应到 $a'x = a'$ 的唯一解 e', 因此单位元素相互对应. 所以, G 中方程 $ax = e$ 的唯一解 a^{-1} 对应到 G' 中方程 $a'x = e'$ 的唯一解 a'^{-1}. 这就完成了证明. 证毕

我们最后证明著名的凯莱 (Cayley) 定理, 这个定理可被解释为是证明变换乘法有关公设的完备性.

定理 8 *任意抽象群 G 与一个变换群同构.*

证明 把由 G 的所有元素组成的 "空间" 上的每个变换 $\phi_a : x \to xa = x\phi_a$ 同 G 的元素 a 联系起来. 因为由 $e\phi_a = e\phi_b$ 可推出 $a = ea = eb = b$, 所以 G 的不同元素对应着不同的变换. 因为对所有的 x, 有

$$x(\phi_a\phi_b) = (x\phi_a)\phi_b = (xa)b = x(ab) = x\phi_{ab}, \tag{6}$$

所以乘积 $\phi_a\phi_b = \phi_{ab}$, 因而所有 ϕ_a 的集合 G' 包含任意两个变换, 就一定包含它们的乘积. 再有, 因为对所有的 x 有 $x\phi_e = xe = x$, 所以 G' 包含单位元素. 我们可以类似地证明, 对所有的 $a, (\phi_a)^{-1}$ 存在, 并在 G' 中, 实际上它就是 ${\phi_a}^{-1}$. 因此 G' 是一个变换群, 根据 (6), 它与 G 同构. 证毕

习　题

1. 下列群中, 任意两个群都同构吗?

   (a) 等边三角形的对称群,　(b) 正方形对称群,
   (c) 正六边形的旋转群,　(d) 模 6 整数加法群.

2. 与习题 1 同样的问题.

   (a) 正方形的旋转群,　(b) 矩形的对称群,
   (c) 菱形 (等边平行四边形) 的对称群,　(d) 模 13 整数 1, 5, 8, 12 的乘法群,
   (e) 模 12 整数 1, 5, 7, 11 的乘法群.

3. (a) 证明: "高斯整数" $m + n\sqrt{-1}(m, n \in \mathbf{Z})$ 的加法群同形为 $2^n3^m(m, n \in \mathbf{Z})$ 的有理分子的乘法群同构.

   (b) 给出两个与矩形网络的变换群同构的群.

*4. 非零实数构成的乘法群与所有实数构成的加法群同构吗？

5. 确定 $\mathbf{Z}_4$ 的加法群与正方形的旋转群之间所有同构.

6. (a) 列出正方形对称群与正方形四个顶点 1, 2, 3, 4 上的变换群之间的同构.

   (b) 像定理 7 那样明显地指出, 两个群中的逆元素在这个同构之下是如何对应的.

7. 对正六边形的旋转群, 做习题 6.

8. 列出与下列每个群同构的变换群, 说明定理 8.

   (a) 所有实数构成的加法群,

   (b) 所有非零实数构成的乘法群,

   (c) 模 8 整数加法群.

6.6 循 环 群

在任意群中, 元素 a 的整数幂a^m 可以分别对正指数、零指数和负指数来定义. 当 $m>0$ 时, 我们定义

$$a^m = a\cdot a\cdots a(m\text{个因子}),\qquad a^0=e,\qquad a^{-m}=(a^{-1})^m. \tag{7}$$

两个普通的指数定律成立:

$$a^ra^s=a^{r+s},\qquad (a^r)^s=a^{rs}. \tag{8}$$

另一方面, 一般来说, $(ab)^r\neq a^rb^r$(参见习题 2).

如果两个指数 r 和 s 都是正的, 那么定律 (8) 可由定义 (7) 直接推出 ①(参见 1.5 节). 对于 (8) 式的第一个定律的其他情形, 当 r 和 s 中有一个可能为零时, (8) 式立即得出; 当 r 和 s 两者都可能为负的时, (8) 式可从定义 (7) 的最后一个公式直接推出. 剩下的情形就是一个指数为负一个指数为正, 比如 $r=-m, s=n$, 其中 $m>0, n>0$. 这时

$$a^{-m}a^n=(a^{-1})^ma^n=\underbrace{(a^{-1}\cdots a^{-1})}_{m\text{个}}\underbrace{(a\cdots a)}_{n\text{个}}.$$

根据结合律我们可以相继消去一些 a 和 a 的逆 a^{-1}. 当 $n\geqslant m$ 时, 留下 a^{n-m}, 而当 $n<m$ 时, 留下某些逆, 即 $(a^{-1})^{m-n}$ 或 $a^{-(m-n)}$. 这两种情形我们都得到所要求的 $a^{-m}a^n=a^{n+(-m)}$.

① r 个因子 "a" 后跟着 s 个因子 "a", 共有 $r+s$ 个因子. 再有, 每组有 r 个因子 "a", s 组共有 rs 个因子.

(8) 式的第二个定律可以更简单地证明. 如果 s 为正, 则由 (8) 式的第一个定律有

$$\underbrace{a^r a^r \cdots a^r}_{s\text{个因子}} = a^{r+r+\cdots+r} = a^{rs}.$$

如果 s 为负, 注意不管 r 是正的、零和负的, 都有 $(a^r)^{-1} = a^{-r}$, 我们可以做类似的展开. 如果 s 为零, 则立即可得结论.

定义 *群中元素 a 的阶是指使得 $a^m = e$ 成立的最小正整数*①*m. 如果找不到 a 的正次幂等于 e, 则定义 a 的阶为无穷. 如果群 G 包含某一个元素 x, G 的元素都由 x 的幂组成, 那么称 G 为循环群; 这个元素 x 称为群 G 的生成元.*

例如, 正方形的所有到自身的旋转构成的群是由 R 的四个幂 R, R^2, R^3 和 $R^4 = I$ 组成, 这里 R 表示顺时针旋转 90°. 这个群完全等同地可以由 R^3 生成 (R^3 表示逆时针旋转 90°), 这是因为 $R^2 = (R^3)^2, R = (R^3)^3, I = (R^3)^4$, 同 R^3 一起组成这个群.

定理 9 *如果元素 a 生成循环群 G, 那么 a 的阶可以确定群 G(在同构意义下). 事实上, 如果 a 的阶是无穷, 那么 G 同构于整数加法群; 如果 a 的阶是某有限整数 n, 那么 G 同构于模 n 整数加法群.*

证明 首先, $a^r = a^s$ 当且仅当

$$e = a^r(a^s)^{-1} = a^r a^{-s} = a^{r-s}, \tag{9}$$

这里用了公式 (8). 再看, 若 $r \neq s$, 则或 $r > s$, 或 $s > r$, 因此, 如果 a 的阶是无穷, 那么不存在整数 $r > s$ 使得 $a^{r-s} = e$, 所以不存在 a 的两个不同的幂是相等的. 此外, 由 (8) 有 $a^s a^t = a^{s+t}$, 因此对应 $a^s \longmapsto s$ 使群 G 与整数加法群同构, 这就证明了定理的第一个结论.

如果 a 的阶是有限的, 那么使 $a^t = e$ 的整数 t 的集合包含零, 由 (8) 可知这个集合还包含它的任意两个元素的和与差. 因此, 根据 1.7 节定理 6, $a^t = e$ 当且仅当 t 是 a 的阶 n 的倍数, 所以根据公式 (9), $a^r = a^s$ 当且仅当 $n|(r-s)$; 也就是说, $a^r = a^s$ 当且仅当 $r \equiv s(\text{mod } n)$. 最后, 再由 (8), 有 $a^r a^s = a^{r+s}$, 所以对应 $a^r \longmapsto r$ 是 G 到模 n 整数加法群的同构. 证毕

定理 9 的一个推论是, 任意循环群 G 的元素个数 (称为群 G 的阶) 等于 G 的任意一个生成元的阶, 任意两个同阶循环群同构.

正方形对称群不是循环群, 不过它是由两个元素 R 和 H 生成的; 事实上, 表 1(6.4 节) 指出

$$R^0 = I, \quad R = R, \quad R^2 = R', \quad R^3 = R'';$$

① 1.4 节的良序原理保证这个 m 一定存在.

$$H = H, \quad HR = D', \quad HR^2 = V, \quad HR^3 = D.$$

于是这个群的全体元素都可唯一地表示成 H^iR^j, 其中 $i = 0, 1$ 与 $j = 0, 1, 2, 3$. 此外, R 和 H 还满足

$$R^4 = I, \qquad H^2 = I, \qquad RH = HR^3.$$

这些等式称为“定义关系”, 因为这些关系可以使任意两个元素的乘积 $H^iR^j (i = 0, 1)$ 化成同样的形式. 例如

$$D'V = HRHR^2 = HHR^3R^2 = IR = R;$$

类似的计算将给出正方形对称群的整个乘法表 (表 1).

习 题

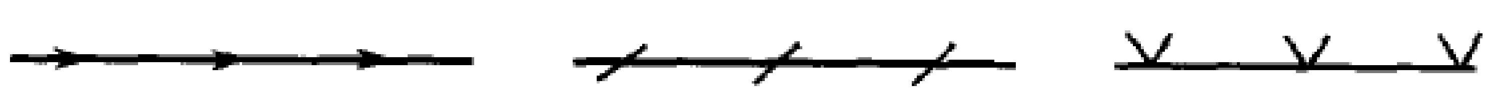
1. 利用定义 $a^1 = a, a^{m+1} = a^m a$, 对正指数用归纳法来证明公式 (8).
2. 证明: 如果对 G 中一切 a, b 及一切正整数 n, 有 $(ab)^n = a^n b^n$, 那么 G 是交换群, 反之亦真.
3. 6 阶循环群有几个不同的生成元?
4. 证明: 如果 6 个元素的交换群包含一个 3 阶元素, 那么这个群是循环群.
5. (a) 模 7 整数 1, 2, $\cdots$, 6 组成的乘法群是循环群吗?

   (b) 模 8 整数 1, 3, 5, 7 组成的乘法群是循环群吗?

   (c) 模 9 整数 1, 2, 4, 5, 7, 8 组成的乘法群是循环群吗?
6. 设循环群 G 是由 m 阶元素 a 生成, 证明: a^k 生成 G 当且仅当 k 与 m 互素.
7. 在习题 6 的假定之下, 求 G 的任意元素 a^k 的阶.
8. 求正方形对称群中每个元素的阶.
9. 给出满足定义关系 $x^2 = y^2 = e, xy = yx$ 的两个元素 x 和 y 生成群的 G 的所有元素和乘法表.
10. 二面体群 D_n 是正 n 边形的所有对称构成的群 (当 $n = 4$ 时, D_n 就是正方形对称群). 证明: D_n 包含 $2n$ 个元素, 并由两个元素 R 和 H 生成, 这里 R 和 H 满足 $R^n = I, H^2 = I, RH = HR^{n-1}$.

*11. 分别找出下面三个无限图案的对称群的生成元和定义关系. 三个群中任意两个同构吗?

想象图形沿两个方向无限延伸下去.



*12. 对 6.3 节中的习题 1、习题 2、习题 4 和习题 5 进行与上题类似的讨论.

6.7　子　群

很多群都包含在较大的群之中. 例如, 正方形的旋转群是正方形对称群的一部分. 再有, 根据对称性诱导出的正方形顶点的八个置换构成的群, 是这些顶点的所有 $4! = 24$ 个置换组成的置换群的一部分. 偶数加法群是整数加法群的一部分.

这些例子提出子群的概念. 群 G 的一个子集 S, 如果关于 G 的二元运算 (乘法) S 本身也是一个群, 那么称 S 为 G 的*子群*.

在任意群 G 中, 仅由单位元素 e 组成的集合是一个子群. 整个群 G 也是它自己的一个子群. G 中不是平凡 ("伪") 子群 e 和 G 的子群称为*真子群*.

定理 10　*群 G 的非空子集 S 是子群当且仅当 (i) 由 a 和 b 在 S 中推出 ab 在 S 中; (ii) 由 a 在 S 中推出 a^{-1} 在 S 中.*

证明　在这些假设之下, 显然 S 是一个子群: 结合律是显然的; 因为至少有一个元素 a 在 S 中, 所以 G 的单位元素 $e = aa^{-1}$ 在 S 中; 群的其他公设已被假定. 反过来, 我们必须证明在任一子群中 (i) 和 (ii) 成立. G 的任意子群的单位元素 $x = e'$ 满足 $xx = x$, 因此它是 G 的单位元素 (6.4 节习题 6). 由此可以推出, 因为对任何元素 a, G 有且仅有一个逆元素, 所以在子群中任何元素 a 的逆元素与它作为 G 中元素的逆元素是同一个元素, 故 (ii) 成立. 条件 (i) 是显然的.　证毕

对于有限阶 (m) 元素 a, 显然有 $a^{m-1}a = a^m = e$, 因此 $a^{-1} = a^{m-1}$. 于是我们有下面简化的条件.

定理 11　*有限群 G 的非空子集 S 是群 G 的子群当且仅当 S 中任意两个元素的乘积仍在 S 中.*

在已知的非阿贝耳群 G 的所有子群中间, 最重要的一个子群是 G 的*中心*. 它定义为, 对一切 $x \in G$ 满足关系 $ax = xa$ 的所有元素 $a \in G$ 的集合. 我们留给读者验证. 实际上, 群的中心总是 G 的子群.

确定一个特定群 G 的全部子群, 一般来说是很困难的. 在 G 是循环群的情况下, 我们现在来确定它的全部子群.

定理 12　*循环群 G 的任何子群是循环群.*

证明　设 G 由元素 a 的幂组成. 如果 a^s 和 a^t 在 S 中, 则由定理 10, $a^{s+t} = a^s a^t$ 和 $a^{s-t} = a^s(a^t)^{-1}$ 都在 S 中. 因此, 使 a^s 在 S 中的整数 s 组成的集合在加法和减法之下是封闭的, 所以 ①这个集合由某一个最小正指数 r 的所有倍数组成 (1.7 节定理 6). 因而 S 由所有幂 $a^{kr} = (a^r)^k$ 组成, 因此 S 是以 a^r 为生成元的循环群.　证毕

在 G 为无限的情况下, 每个 $r > 0$ 确定不同的子群. 如果 G 有 n 个元素, 那么, 因为 $a^n = e$ 一定在 S 中, 所以只有那些 n 的因子 $r > 0$ 才能用这种方法确定出 G

① 当集合 S 仅由零组成时, 取 $r = 0$, 这个结论还成立.

的子群, 而这些子群全不相同.

为得到进一步研究子群的材料, 我们现在列出正方形对称群的全部子群. 通过验证 6.1 节给出的这个群运算的定义, 我们找到了全部真子群, 每个子群保持了下列八种构形中的一种不变性:

一对角线	一轴	一面
$[I, D, D', R']$	$[I, H, V, R']$	$[I, R, R', R'']$
一轴和一对角线	顶点 1(或 3)	顶点 2(或 4)
$[I, R']$	$[I, D]$	$[I, D']$
一垂直边	一水平边	
$[I, H]$	$[I, V]$	

这里保持面不变的变换, 我们理解为正方形没有翻转的那些变换. 所有这些子群可以在一个表上按它们相互之间的关系表示出来, 其中每个群用向下的线或一串线同它的所有子群连接起来, 如图 6-3 所示.

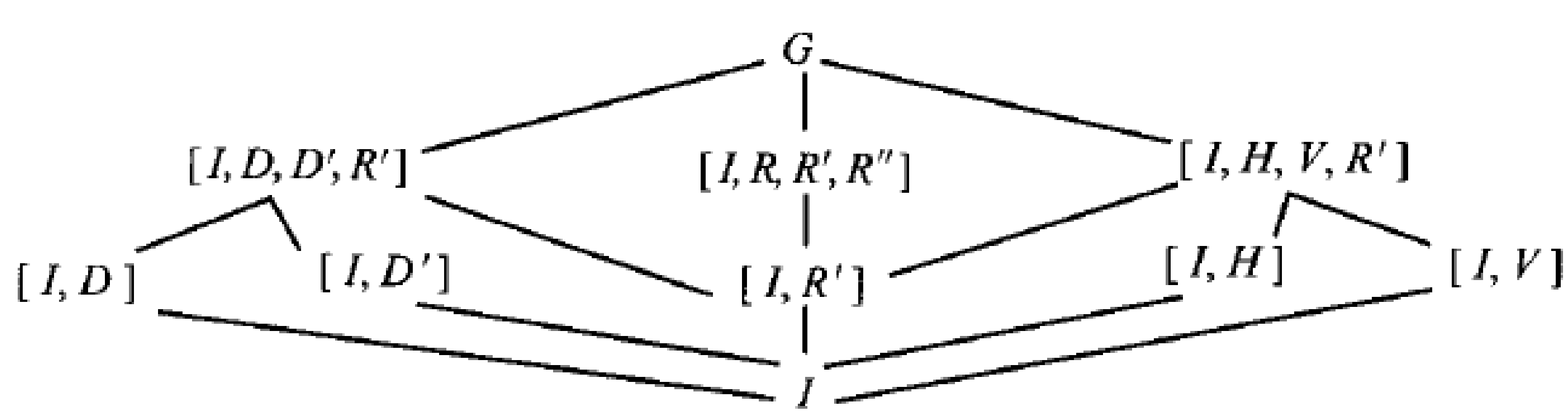


图　6-3

不用几何方法我们也可以找出所有这些子群. 事实上, 把群的元素看作纯抽象的元素, 可以最有效地确定一个特定有限群 G 的全部子群, 如下所述.

首先注意, 如果 G 的子群 S 包含元素 a, 则 S 也包含由 a 的所有幂组成的循环子群 $\{a\}$(证明它是子群!). 在上述例子中, 这个方法给出列出的除前两个子群以外的所有子群. 其次注意, 任何子群不仅包含两个循环子群 $\{a\}$ 和 $\{b\}$, 而且必包含 a 和 b 的幂的所有乘积 (例如 $a^2b^{-3}a$) 所组成的集合 $\{a, b\}$. (用定理 11 证明这个集合构成一个子群!) 在上述例子中, 这种方法给出了剩下的子群. (在 6.8 节我们将看到, 为什么这些子群都是含有 2 个或 4 个元素.) 一般来说, 我们还可以进一步对由三个或更多的元素生成的子群 $\{a, b, c\}$ 进行检验, 但是这时群中元素的个数应至少是四个不同素数之积, 否则决不会发生这种情况.

两个子群 (实际上也可以是任意两个集合!)S 和 T 的交$S \cap T$ 是由既属于 S 又属于 T 的所有元素组成的集合.

定理 13　*群 G 中两个子群 S 和 T 的交 $S \cap T$ 是 G 的子群.*

证明　根据定理 10, a 在 $S \cap T$ 中意味着 a 在 S 中, 因此 a^{-1} 在 S 中; 同样可推出

a^{-1} 在 T 中, 所以 a^{-1} 在 $S\cap T$ 中. 类似地, a 和 b 都在 $S\cap T$ 中意味着 ab 既在 S 中又在 T 中, 所以 ab 在 $S\cap T$ 中. 因此根据定理 10, $S\cap T$ 是一个子群. 还有, $S\cap T$ 包含 e, 所以 $S\cap T$ 是非空的.　证毕

显然, $S\cap T$ 是包含在 S 和 T 中的最大子群. 对偶地, 存在包含 S 和 T 的最小子群. 它由 S 和 T 中元素的正幂和负幂的所有乘积组成, 称它为 S 和 T 的并, 记作 $S\cup T$. 在第 11 章中我们将再讨论这些概念.

习　题

1. 在正六边形对称群中, 保持对角线不变的子群是什么?
2. 证明: 如果 T 是 S 的子群, S 又是 G 的子群, 那么 T 是 G 的子群.
3. 在四个数字 1, 2, 3, 4 的置换群中 (置换记作 ϕ), 找出下列子群:
   (a) 所有把集合 $\{1,2\}$ 变为 $\{1,2\}$ 的置换 ϕ;
   (b) 所有适合 "对集合 $\{1,2,3,4\}$ 中任意两个数字 a,b, 由 $a\equiv b(\mathrm{mod}\ 2)$ 可推出 $a\phi\equiv b\phi(\mathrm{mod}\ 2)$" 的置换 ϕ.
4. 证明: 当 G 是无限的, 但 G 的所有元素有有限阶, 定理 11 仍然成立. 说明 $\mathbf{Z}_p[x]$ 的加法群就是这样一个群.
5. 列出下列各群的所有子群:
   (a) 模 12 整数加法群;　(b) 正五边形对称群;
   (c) 正六边形对称群;　*(d) 四个字母的置换群.

*6. 设 $a\leftrightarrow a'$ 是两个置换群 G 和 G' 之间的同构, 又设 S 是 G 中保留一个字母固定的那些置换组成的集合. G' 中与所有 $a\in S$ 对应的那些元素组成的集合 S', 一定是 G' 的子群吗? 集合 S' 一定保留一个字母固定吗? 说明一下.

7. 证明: 任意群 G 的中心是 G 的子群.
8. 找出下列各群的中心:
   (a) 正方形对称群;　(b) 等边三角形对称群.

*9. 找出正 n 边形对称群的中心.

*10. 证明: 任意交换群 G 中的全体有限阶元素构成 G 的一个子群.

6.8　拉格朗日定理

我们现在来讨论抽象群理论中一具有重要意义的概念: 群 G 的任意子群 S 分解 G 成陪集.

定义　*群或子群的阶指的是它的元素个数. 设 S 是群 G 的一个子群, a 是 G 中一个固定元素, 则 S 的所有元素 s 用 a 右 (左) 乘的右 (左) 倍数 $sa(as)$ 所有组成的集*

合 $Sa(aS)$ 称为 G 的子群 S 在 G 中的一个右 (左) 陪集. S 的不同右陪集的个数称为子群S 在 G 中的“指数”.

因为 $Se = S$, 所以 S 是它本身的一个右陪集. 此外我们有

引理 1 *如果 S 是有限的, 则 S 的每个右陪集 Sa 中元素的个数同 S 的元素一样多.*

这是因为, 变换 $s \longmapsto sa$ 是双射: 右陪集 Sa 的每个元素 $t = sa$ 是 S 的元素 $s = ta^{-1}$ 的像, 这个元素是唯一的. (参见定理 8).

引理 2 *S 的两个右陪集 Sa 和 Sb, 或者相等, 或者没有公共元素.*

这是因为, 假定 Sa 和 Sb 有一个公共元素 $c = s'a = s''b(s', s''$在 S 中). 那么 Sb 包含 Sa 的每个元素 $sa = ss'^{-1}s'a = ss'^{-1}s''b = (ss'^{-1}s'')b$. 类似地, Sa 包含 Sb 的每个元素. 所以 $Sa = Sb$.

举例说明这些引理是容易的. 例如, 如果 G 是正方形对称群, 则子群 $S = [I, H]$ 有四个右陪集:

$$[I,H]I = [I,H], \qquad [I,H]R = [R,HR] = [R,D'],$$
$$[I,H]R' = [R',HR'] = [R',V], \qquad [I,H]R'' = [R'',HR''] = [R'',D].$$

每个陪集有两个元素, 并且对称群中的每个元素都落入这四个右陪集中的一个.

再有, 如果 G 是整数加法群, 则由 5 的倍数 $\pm 5n$ 组成的子群, 它的所有右陪集就是模 5 的不同剩余类. 最后, 设 G 是数字 $1, 2, \cdots, 6$ 的所有置换组成的对称群, 而 S 是保持数字 1 固定的置换组成的子群. 那么由 $1\phi = k$ 可推出, 对所有的 $\psi \in S$, 有 $1(\psi\phi) = (1\psi)\phi = 1\phi = k$. 因此陪集 $S\phi$ 只包含 5! 个把 1 变为 k 的置换 (根据引理 1, 这是 $S\phi$ 的全部元素). 所以 S 的右陪集是 G 中分别使 $1 \longmapsto 1, 1 \longmapsto 2, \cdots, 1 \longmapsto 6$ 的子集合.

从上述这些引理我们得到一个经典的结果, 这个结果对有限群的理论来说是基本的和重要的. 因为任意右陪集 Sa 总包含 $a = ea$, 所以任意群 G 的每个元素都包含在某一个右陪集中. 因此 G 可用 S 分解成一些不重迭的子集合, 每个子集合的元素恰恰同 S 的元素一样多. 如果 G 是有限的, [①] 这个结论就是

定理 14 (拉格朗日) *有限群 G 的阶是它的每个子群的阶的倍数.*

G 的每个元素 a 生成一个循环子群, 它的阶就是 a 的阶 (定理 9 的推论). 因此我们有

推论 1 *有限群 G 的每个元素的阶都是 G 的阶的因子.*

推论 2 *具有素数阶 p 的群是循环群.*

这是因为, 在有限群中, 由任意元素 $a \neq e$ 生成的循环子群 A 的阶 $n > 1$, 可整除 p. 而这就意味着 $n = p$, 因此 $G = A$ 是循环群.

---

① 推广到无限的情况, 可从第 12 章的讨论中立即得到, 但这并不重要.

更一般地, 拉格朗日定理可以用来确定 (精确到同构) 所有任意低阶的抽象群. 例如, *四群*是定义为由四个可交换元素: e(单位元素) 和 $a, b, c = ab$ 组成的群, 后面三个元素的阶都是 2. 在 6.9 节中我们将证明这个群与矩形对称群同构. 我们现在证明

推论 3 *四阶抽象群只有四阶循环群和四群两种.*

换句话说, 每个四阶群或者同构于四阶循环群, 或者同构于四群.

证明 当四阶群包含一个四阶元素时, 这个群是循环群. 否则, 由推论 1 知, G 的元素除 e 外, 它们的阶一定都是 2. 记它们为 a, b, c. 根据消去律, ab 不可能是 $ae = a, eb = b$ 或 $aa = e$, 因此 $ab = c$. 类似地, $ba = c, ac = ca = b, bc = cb = a$. 而这些等式连同 $a^2 = b^2 = c^2 = e$ 和对一切 x 有 $ex = xe = x$ 一起给出四群的乘法表.

拉格朗日定理也可以应用到数论中.

推论 4(费马) *如果 a 是整数, p 是素数, 那么 $a^p \equiv a(\text{mod}\, p)$.*

证明 模 p 整数 (零除外) 乘法群有 $p-1$ 个元素. 那么根据推论 1, 这个群的任意元素 a 的阶是 $p-1$ 的因子, 所以对任何元素 $a \not\equiv 0\ (\text{mod}\, p)$ 有 $a^{p-1} \equiv 1(\text{mod}\, p)$. 如果我们用 a 乘同余式两边, 我们就得到所要求的同余式. 对于 $a \equiv 0(\text{mod}\, p)$ 的情况, 结论显然正确.(这是第 1 章定理 18 的一个新证法.)

习 题

1. 对 $p = 7$ 和 $a = 2, 3, 6$ 验证费马定理.
2. (a) 列出 26 阶二面体群 (6.6 节习题 10) 的全部子群. 共有多少个子群?
   (b) 推广你的结果.
3. 证明: 有限群的任意子群的右陪集的个数等于它的左陪集的个数. (*提示*: 利用对应 $x \longmapsto x^{-1}$.)
4. 确定正方形对称群的子群 $[I, D]$ 的陪集.
5. 设 S 是群 G 的任意子群, 又设 SaS 表示由所有乘积 sas'(s, s'在S中) 组成的集合. 证明: 对任意 $a, b \in G$, 或者 $SaS \bigcap SbS$ 是空集, 或者 $SaS = SbS$.
6. 对任意子群 S, 设 $x \equiv y(\text{mod} S)$ 是指 $xy^{-1} \in S$.
   (a) 证明: 这个关系满足自反律、对称律和传递律. 并证明: $x \equiv y(\text{mod}\, S)$ 当且仅当 x 和 y 在 S 的同一个右陪集中.
   (b) 证明: 由 $x \equiv y(\text{mod}\, S)$ 可推出, 对一切 a 有 $xa \equiv ya(\text{mod}\, S)$.
7. 设 G 是正六边形对称群, S 是保持一个顶点固定的子群. 求出 S 的右陪集和左陪集.
8. 证明: p^m 阶群 (这里 p 为素数) 一定包含一个 p 阶子群.
9. (a) 设 G 是 $\mathbf{R}$ 上所有变换 $x \longmapsto ax + b$(其中 $a \neq 0$, b 为实数) 构成的群, 而 S 是 $a = 1$ 的所有这样的变换构成的子群. 描述 S 在 G 中的右陪集和左陪集.
   (b) 又设 T 是 $b = 0$ 的所有这样的变换构成的子群, 描述 T 在 G 中的右陪集和左陪集.

*10. (a) 证明：在任意交换环 R 中, 所有单位 (具有乘法逆元素的那些元素) 构成一个群 G.

(b) 证明：如果 $R = \mathbf{Z}_n$, 那么 G 是由所有与 n 互素的正整数 $k < n$ 组成.

(c) 在 $R = \mathbf{Z}_n$ 的情况下, G 的阶记作 $\phi(n)$, 并称为欧拉函数. 证明：当 $n = p$ 为素数时, $\phi(p) = p - 1$. 计算 $\phi(12), \phi(16), \phi(30)$.

(d) 用拉格朗日定理证明：如果 $(k, n) = 1$, 那么 $k^{\phi(n)} \equiv 1 (\mathrm{mod}\ n)$.

*11. 证明：如果 S 和 T 分别是群 G 的 s 阶和 t 阶子群, 并且 $S \bigcap T$ 和 $S \bigcup T$ 的阶分别为 u 和 v, 那么 $st \leqslant uv$.

*12. 证明：6 阶抽象群只有 6 阶循环群和三字母的对称群.

*13. 设 $2^h + 1$ 是素数 p.

(a) 证明：模 p 整数乘法群中, 2 的阶是 $2h$.

(b) 利用费马定理推证：$2h$ 可整除 $p - 1 = 2^h$.

(c) 导出结论 h 是 2 的幂.

6.9 置 换 群

置换是有限集合到自身的一一变换. 例如, 由 1, 2, 3, 4, 5 五个数字可以组成一个集合. 一个置换可以是变换 ϕ：

$$1\phi = 2, \quad 2\phi = 3, \quad 3\phi = 4, \quad 4\phi = 5, \quad 5\phi = 1. \tag{10}$$

另一个置换可以是变换 ϕ'：

$$1\phi' = 2, \quad 2\phi' = 3, \quad 3\phi' = 1, \quad 4\phi' = 5, \quad 5\phi' = 4. \tag{11}$$

读者会发现, 计算出 $\phi\phi'$ 和 $\phi'\phi$, 并注意 $\phi\phi' \neq \phi'\phi$ 是有益的.

一个置换, 像上面定义的置换 ϕ 那样, 如果它给出置换符号的一个循环排列 (图 6-4), 那么这个置换称为*循环置换*或称为*循环*. 为表示循环置换, 有一个含蓄的记号 —— 仅仅把字母写到括号里面, 首先写出所包含的任何一个字母, 然后写出它变换后的字母,……, 最后写出能变换成原来第一个字母的那个字母. 例如, (10) 式表示的置换 ϕ 可以写成下列等价形式中的任何一个：

$$(12345), (23451), (34512), (45123), (51234).$$

定理 15 *n 个符号的循环置换的阶是 n.*

证明 循环置换 $\gamma = (a_1 a_2 \cdots a_n)$ 把 a_i 变成 a_{i+1}. 因此 γ^2 的效果相当于 γ 作用两次, 把每个 a_i 变成 a_{i+2}. 一般地, γ^k 把 a_i 变成 a_{i+k}, 这里所有下标都按模 n 化简了.

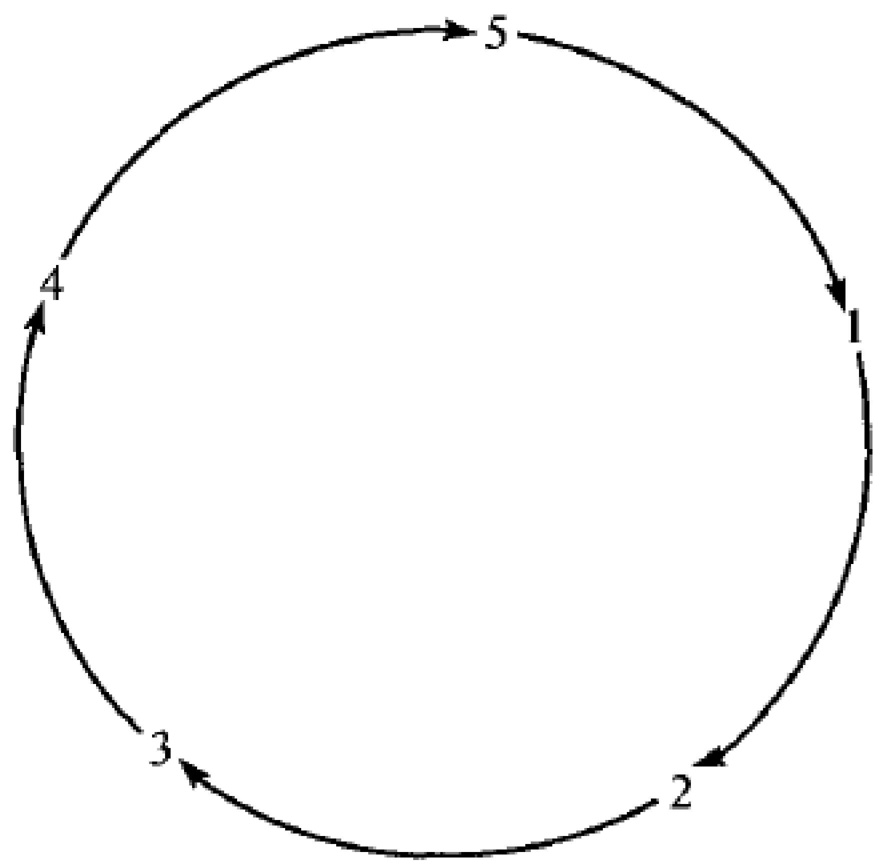


图 6-4

γ^k 为单位元素 I 当且仅当 a_{i+k} 等于 a_i, 即当且仅当 $k \equiv 0 (\bmod n)$. 因为使得 $\gamma^k = I$ 的最小整数 k 是 n 本身, 所以 γ 的阶是 n(见 6.6 节中的定义). 这时我们说循环 γ 的长度是 n.

循环置换的记号可以推广到任意置换的情形. 例如, (11) 式中表示的置换 ϕ' 把数字 1, 2 和 3 循环排列, 并且把 4 和 5 循环排列. 于是 ϕ' 是这两个循环的积

$$(123)(45) = (45)(123).$$

这个乘积可以按两种次序写, 是因为由 (123) 置换过的符号在 (45) 作用下保持不变, 这表示按两种次序相继使用这两个置换, 其结果一样.

定理 16 任意置换 ϕ 可写成几个循环的乘积, 这些循环分别作用在不相交①的符号集上 (更简洁些说, 任意置换 ϕ 可写成几个不相交的循环之积).

证明 选择任意一个符号记作 a_1, 再用 a_2 表示 $a_1\phi$, 用 a_3 表示 $a_2\phi, \cdots\cdots$, 用 a_n 表示 $a_{n-1}\phi$, 直到 $a_n\phi = a_i$ 是前面某一个已经命名了的元素. 因为任意 $a_i (i > 1)$ 前面一个元素是 a_{i-1}, 所以 $a_n\phi$ 一定是 a_1. 于是 ϕ 作用到字母 $a_1, a_2, \cdots, a_n$ 上的结果是循环 $(a_1a_2\cdots a_n)$. 此外, 当循环 $(a_1a_2\cdots a_n)$ 包含任意字母 a_i 时就一定包含前一个字母 a_{i-1}, 因此 ϕ 还要置换除这 n 个字母外剩下来的字母. 现在对符号的个数用归纳法就可推出定理的结论. 特别地, m 个字母的恒等置换可表示成 m 个循环之积, 每个循环的长度为 1.

反之, 显然任意不相交循环之积是一个置换. 此外我们可以证明

定理 17 任意置换 ϕ 的阶等于 ϕ 的不相交循环之长度的最小公倍数.

证明 把置换 ϕ 写成不相交循环 $\gamma_1, \cdots, \gamma_r$ 的乘积 $\phi = \gamma_1 \cdots \gamma_r$. 如果 $i \neq j$ 则 γ_i 和 γ_j 是不相交的; 因此 $\gamma_i\gamma_j = \gamma_j\gamma_i$, 并且因子 γ_i 可以在 ϕ 和它的幂中重新排列, 从而

① 两个集合不相交是指它们没有公共元素.

对所有整数 n, 得到 $\phi^n = \gamma_1^n \cdots \gamma_r^n$. 所以 $\phi^n = I$ 当且仅当每个 γ_i^n 是恒等置换. 而根据定理 15, 由此可推出, $\phi^n = I$ 当且仅当 n 是 $\gamma_1, \cdots, \gamma_r$ 的长度的公倍数, 由此立即得到定理 17 的结论. 证毕

根据 6.5 节的定理 8, 每个有限群同构于一个或多个置换群. 特别地, 这对于由几何图形的对称构成的有限群是正确的, 我们现在用两个例子来说明这一点.

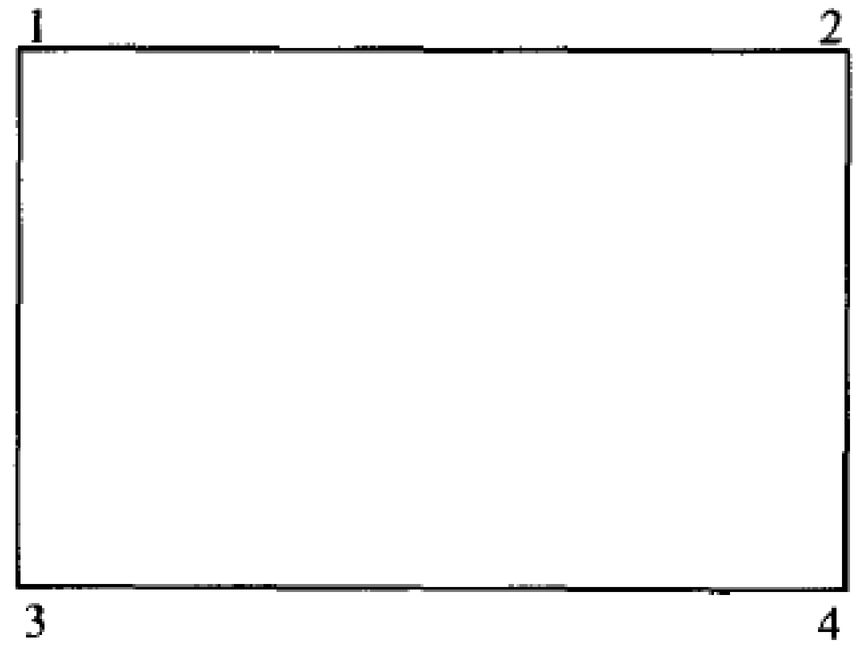


图 6-5

考虑矩形对称群 (图 6-5). 在这个群中, 由下列四个置换

$$I = (1)(2)(3)(4), \quad R = (14)(23), \quad H = (13)(24), \quad V = (12)(34).$$

来变换它的顶点. 这个群被称为*四群*. 根据定理 8, 它同构于置换群

$$\phi_I = (I)(R)(V)(H), \qquad \phi_R = (IR)(HV),$$

$$\phi_H = (IH)(RV), \qquad \phi_V = (IV)(RH).$$

类似地, 正方形对称群 (6.1 节) 可以表示为四个顶点的置换群. 利用定理 8, 我们也可以把它表示为八个符号的置换群, 其中每个符号代表正方形对称群的一个元素. 例如, 符号 R 对应于一个置换, 这个置换的效果是这八个符号用 R 右乘后所得到的元素, 我们从正方形对称群表 (表 1) 中以 R 为首的那一列看出, 这个置换就是 $(IRR'R'')(HD'VD)$. 类似地, 符号 H 对应于置换 $(IH)(RD)(R'V)(R''D')$.

相同长度的两个循环有着密切的关系. 例如, 如果 $\gamma = (1234)$ 和 $\gamma' = (2143)$, 那么我们可以计算出 $\gamma' = \phi^{-1}\gamma\phi$, 其中 $\phi = (12)(34)$ 是一个置换, 它把循环 γ 中每个数字变换成 γ' 中相应的数字. 这是下面结果的特殊情形.

定理 18 *设 ϕ 和 γ 是 m 个字母的置换, 其中 γ 是循环置换 $\gamma = (a_1 \cdots a_m)$, 并用 $\gamma' = (a_1\phi \cdots a_m\phi)$ 表示另一个循环, 它是用 γ 表示式中每个 a_i 在 ϕ 作用下所成的像 $a_i\phi$ 来代替 a_i 而得到的. 那么 $\phi^{-1}\gamma\phi = \gamma'$.*

证明 乘积 $\phi^{-1}\gamma\phi$ 把每个字母 $a_i\phi$ 先映成 $a_i\phi\phi^{-1} = a_i$, 再映成 $a_i\gamma = a_{i+1}$, 最后映成 $a_{i+1}\phi$, 因此 $\phi^{-1}\gamma\phi$ 作用到 $a_i\phi$ 上的效果与 γ' 作用到 $a_i\phi$ 上的效果相同 (记

$a_{m+1}=a_1$). 类似地, 我们计算出, $\phi^{-1}\gamma\phi$ 和 γ' 都把任意不是形为 $a_i\phi$ 的字母 b 变为自身. 因此 $\phi^{-1}\gamma\phi=\gamma'$, 如断言所述.

推论 对任意两个置换 ϕ 和 ψ, 如果 ψ 写成循环之积 $\psi=\gamma_1\cdots\gamma_r$, 那么我们有

$$\phi^{-1}\psi\phi=\gamma_1'\cdots\gamma_r',$$

式中 γ_i' 是从 γ_i 得到的, 如定理 18 所述.

习　题

1. 把下列置换 ϕ 表示成不相交循环之积:

   (a) $1\phi=4$, $2\phi=6$, $3\phi=5$, $4\phi=1$, $5\phi=3$, $6\phi=2$;

   (b) $1\phi=5$, $2\phi=3$, $3\phi=2$, $4\phi=6$, $5\phi=4$, $6\phi=1$;

   (c) $1\phi=3$, $2\phi=5$, $3\phi=6$, $4\phi=4$, $5\phi=1$, $6\phi=2$.

   求出每个置换的阶.

2. 把下列乘积表示成不相交循环之积:

   (1234) (567) (261) (47),　(12345) (67) (1357) (163),　(14) (123) (45) (14).

   求出每个乘积的阶.

3. 求置换 $(abcdef)(ghij)(klm)$和$(abcdef)(abcd)(abc)$ 的阶.

4. 把菱形对称群表示成它的四个顶点的置换群.

5. 描述由所有那些把 $\{x_1,x_2\}$ 集合映到自身的 $x_1,\cdots,x_6$ 的置换构成的子群的所有右陪集和左陪集.

6. 哪些对称群是阿贝耳群?

7. 设 G 是由保持一个顶点固定的立方体所有对称构成的群, 把 G 表示成这些顶点的置换群 (参见 6.3 节).

8. (a) 证明: 每个置换可写成长度为 2 的循环 (“对换”) 之积 (一般来说, 不一定不相交).

   *(b) 这个结论同“由 $ab=ba$ 证明一般交换律” (1.5 节) 有什么关系?

9. 把等边三角形对称群表示成

   (a) 三个字母的置换群;　　(b) 六个字母的置换群;

   *(c) 用两种本质上不同的方式做 (b).

*10. 证明: n 次对称群是由循环 $(12\cdots(n-1))$ 和 $((n-1)n)$ 生成的.

*11. 在什么意义下, 定理 16 的表达式是唯一的? 证明你的回答.

6.10　偶置换与奇置换

当我们考虑齐次多项式形式

$$P = \prod_{i<j}(x_i - x_j) \qquad (\text{其中 } i, j \text{ 从 1 跑到 } n)$$

时, 就会发现置换的一个重要分类. 当 $n = 3$ 时,

$$\begin{aligned} P &= (x_1 - x_2)(x_1 - x_3)(x_2 - x_3) \\ &= x_1^2x_2 + x_2^2x_3 + x_3^2x_1 - x_1^2x_3 - x_3^2x_2 - x_2^2x_1, \end{aligned}$$

而且 P^2 是 5.5 节中讨论的判别式. 一般地, P 是 $\frac{n(n-1)}{2}$ 次多项式. 显然, P 中下标的任意置换使 P 的这组因子保持不变, 因此除符号外, P 本身也不变. 而且对换 (x_1x_2) 把 $(x_1 - x_2)$ 变为它的负值 $(x_2 - x_1)$, 把 $(x_1 - x_j)$ 和 $(x_2 - x_j)$ 互换 $(j > 2)$, 并保持其他因子不变, 因此 (x_1x_2) 把 P 变为 $-P$.

因此全体下标的 $n!$ 个置换分为两类: 保持 P(或 $-P$) 不变的置换称为*偶置换*, 把 P 和 $-P$ 互换的置换称为*奇置换*. 由此推出, 当我们考虑相继实行这两类置换的效果时, 有下列法则

$$\begin{aligned} &\text{偶置换} \times \text{偶置换} = \text{奇置换} \times \text{奇置换} = \text{偶置换}, \\ &\text{偶置换} \times \text{奇置换} = \text{奇置换} \times \text{偶置换} = \text{奇置换}. \end{aligned} \tag{12}$$

我们用公式 (12) 和定理 11 得出一个推论: 全体偶置换构成 n 次对称群的子群 A_n. 这个子群通常称为 n 阶 "交错群".

此外, 如果 β 是固定的奇置换, ϕ 是变化的奇置换, 那么 $\phi\beta^{-1}$ 是偶置换, 所以 $\phi = (\phi\beta^{-1})\beta$ 是在右陪集 $A_n\beta$ 中. 概括地说, 全体奇置换构成 A_n 的单个右陪集. 因此根据拉格朗日定理, n 个符号的 "交错群" 刚好包含 $\frac{n!}{2}$ 个元素.

n 个未定元 $x_1, \cdots, x_n$ 的多项式 $g(x_1, \cdots, x_n)$, 如果在它的下标置换的对称群作用下它是不变的, 则称 $g(x_1, \cdots, x_n)$ 为 "对称" 多项式. 特殊对称多项式是 (对 $n = 3$)

$$\sigma_1 = x_1 + x_2 + x_3, \quad \sigma_2 = x_1x_2 + x_1x_3 + x_2x_3, \quad \sigma_3 = x_1x_2x_3. \tag{13}$$

它们是下面展开式的系数

$$(t - x_1)(t - x_2)(t - x_3) = t^3 - \sigma_1t^2 + \sigma_2t - \sigma_3. \tag{14}$$

一般地, 我们称这样的多项式为*初等对称多项式*(n 个变量), 它们是

$$\sigma_1 = \sum_i x_i, \qquad \sigma_2 = \sum_{i<j} x_ix_j, \sigma_3 = \sum_{i<j<k} x_ix_jx_k, \ \cdots, \ \sigma_n = x_1 \cdots x_n. \tag{15}$$

因为 $(-1)^k\sigma_k$ 是 $p(t)=\prod_k(t-x_k)$ 作为 t 的多项式的展开式中 t^{n-k} 的系数, 这些表达式 σ_i 给出 $p(t)$ 的系数, 这些系数为 $p(t)$ 的根的函数. 从所谓"对称多项式基本定理"可推导出初等对称多项式的很多重要性质. 现在我们不加证明地[①]叙述这个定理.

定理 19 *任意对称多项式 $p(x_1,\cdots,x_n)$ 可表示为初等对称多项式的多项式.*

例如, 在两个变量 x 和 y 的情况下,

$$x^2+y^2=(x+y)^2-2xy=\sigma_1^2-2\sigma_2,$$
$$x^3+y^3=(x+y)^3-3xy(x+y)=\sigma_1(\sigma_1^2-3\sigma_2), \text{等等}.$$

即使多项式 $q(x_1,\cdots,x_n)$ 不是对称的, 我们至少也可以要求找出保持多项式不变的所有那些下标置换组成的集合. 显然这个集合是一个群; 它称为多项式群.

习 题

1. (a) 列出三个字母的奇置换. (b) 列出四个字母的奇置换.
2. 对哪些正整数 n, 长度为 n 的循环是偶置换? 对哪些正整数 n , 长度为 n 的循环是奇置换?
3. (a) 证明: 若干个循环 (不一定不相交) 的乘积是奇置换当且仅当它包含着奇数个长度为偶数的循环.

   (b) 置换 (123)(246)(5432) 和 (12)(345)(67)(891) 是奇置换还是偶置换?
4. (a) 构造 11 个字母的 14 阶偶置换和奇置换的例子.

   (b) 证明每个 8 个字母的 10 阶置换是奇置换.
5. 证明: 一个置换是偶置换当且仅当它可以写成偶数个对换 (6.9 节习题 8) 之积.

*6. 证明: 每个偶置换可以写成长度为 3 的循环之积.

7. 求出下列每个多项式的多项式群:

$$x_1x_2+x_3x_4,\quad x_2x_1+x_3x_2+x_2x_4,\quad x_1^2x_2+x_3x_4^2+x_1^2x_3+x_2x_4^2.$$

8. 用初等对称多项式表示下列多项式:

$$x^2+y^2+z^2,\qquad x^2y+y^2z+z^2x+x^2z+y^2x+z^2y.$$

6.11 同 态

从群 G 到群 G' 的单值变换可以保持乘法, 但不是一一的 (也就是说, 这个变换

① 参见 L. Weisner, *Introduction to the Theory of Equations* (New York: Macmillan, 1938), p 108. 也可参见 15.6 节定理 15 的推论.

不是同构).

例如, 考虑 n 次对称群和 ± 1 在乘法之下构成的群之间的对应, 这个对应把偶置换映射到 $+1$, 把奇置换映射到 -1, 由公式 (12), 它把乘积映射到乘积.

或者考虑对应 $n \longmapsto \mathrm{i}^n$, 其中 $\mathrm{i}=\sqrt{-1}$, 这是整数加法群和四次单位根的乘法群之间的对应. 它也保持了群的运算: $\mathrm{i}^{m+n}=\mathrm{i}^m\mathrm{i}^n$, 但这个对应是多对一的.

这些例子和其他的例子引出下面的概念.

定义 *把群 G 映射到群 G' 的单值变换 $x \longmapsto x'$, 如果对 G 中一切 x, y 有 $(xy)'=x'y'$, 则称这个变换为群 G 到群 G' 的同态.*

定理 20 *在任意同态 $G \to G'$ 之下, G 的单位元素映射到 G' 的单位元素, G 的逆元素映射到 G' 的逆元素.*

证明 因为 $e^2=e$, 所以 e 的像 f 满足 $f^2=f=fe'$, 这时 e' 是 G' 的单位元素. 因此根据消去律有 $f=e'$, 所以 G 的单位元素一定映射到 G' 的单位元素. 同样, 如果 a 映射到 a', a^{-1} 映射到 $(a^{-1})'$, 那么 $aa^{-1}=e$ 一定映射到 $a'(a^{-1})'=e'$, 所以 $(a^{-1})'$ 是 a' 的逆元素.

推论 1 *循环群的任意同态像*[①] *是循环群.*

因为根据定理 20, 不论 m 是正整数、零或负整数, 都有 $(a^m)'=(a')^m$, 因此, 如果群 G 由所有幂 a^m 构成, 那么群 G' 也由 a' 的所有幂 $(a')^m=(a^m)'$ 构成.

推论 2 *在群 G 到群 G' 的同态之下, G 中映射到 G' 的单位元素 e' 的所有元素组成的集合 N 是 G 的一个子群.*

这个集合 N 称为这个同态的核.

因为 $e \longmapsto e'$, 所以 N 是非空的. 再根据定理 20 和假设条件, 由 $a \longmapsto e'$ 和 $b \longmapsto e'$ 可推出 $a^{-1} \longmapsto (a')^{-1}=(e')^{-1}=e'$ 和 $ab \longmapsto a'b'=e'e'=e'$, 因此 N 是一个子群.

直积 任意两个群 G 和 H 有*直积* $G \times H$, $G \times H$ 的元素都是有序对 (g,h), 其中 $g \in G, h \in H$; $G \times H$ 中的乘法由下面公式定义:

$$(g,h)(g',h')=(gg',hh'). \tag{15a}$$

显然, (e,e) 在 $G \times H$ 中起单位元素的作用; (g^{-1},h^{-1}) 是 (g,h) 的逆元素; 并且乘法满足结合律; 因此 $G \times H$ 是一个群. 此外, 函数 $\alpha(g,h)=g$ 定义一个从 $G \times H$ 到 G 上的同态 α, 函数 $\beta(g,h)=h$ 定义一个从 $G \times H$ 到 H 上的同态 β.

还可以证明, 每个有限阶的阿贝耳群与阶为素数幂循环群的直积同构. 我们这里只需用下面比它更弱的结果.

定理 21 *设 m 和 n 互素, 则 m 阶循环群和 n 阶循环群的直积是一个 mn 阶循环群.*

① 同态映上有时称为满同态, 于是相应地称同态像为满同态像.

证明 设 a 和 b 分别生成循环群 A 和 B, 其阶分别为 m 和 n. 那么, 在 $C = A \times B$ 中, $(a,b)^k = (a^k, b^k)$ 是单位元素 (e,e) 当且仅当 $k \equiv 0 (\mathrm{mod}\ m), k \equiv 0 (\mathrm{mod}\ n)$. 根据 1.9 节定理 17, 这就意味着 $k \equiv 0 (\mathrm{mod}\ mn)$. 因此 $(a,b) = c$ 是 C 中 mn 阶元素, C 只包含 mn 个元素, 所以它是循环群. 证毕

习 题

1. 在同态 $n \longmapsto \mathrm{i}^n$(其中 $\mathrm{i} = \sqrt{-1}, n \in \mathbf{Z}$) 下, 求出同态核.
2. 证明: 8 阶循环群有同态像为

   (a) 4 阶循环群, (b) 2 阶循环群.
3. 把每个 x 映射到复数 $\mathrm{e}^{2\pi \mathrm{i} x}$ 上的对应是所有实数 x 构成的加法群的同态吗? 如果是, 那么它的同态像是什么? 它的同态核是什么?
4. 设 G 是 n 个字母 $1, 2, \cdots, n$ 的某些置换组成的群, G 中每个置换 ϕ 把字母 $1, 2, \cdots, k$ 的子集合映射到自身, 又设 G' 是字母 $1, 2, \cdots, k$ 的全体置换 ϕ^* 组成的群, 证明群 G 满同态于群 G'.
5. 在正方形中, 设两条对角线为 d 和 d', 两个轴是 h 和 v. 证明: 存在一个同态 $\phi \longmapsto \phi^*$, 在这个同态下, 正方形对称群的每个运动 ϕ 诱导出关于 d, d', h 和 v 的置换 ϕ^*. 详细列出对应 $\phi \longmapsto \phi^*$. 它的同态核是什么?
6. 证明: 如果 G 同态于 G', G' 同态于 G'', 那么 G 同态于 G''.
7. 下列这些对应中, 哪些是所有非零实数的乘法群到自身的同态? 如果对应是同态, 指出它的同态像 G' 和同态核.

   (a) $x \longmapsto |x|$, (b) $x \longmapsto 2x$, (c) $x \longmapsto x^2$, (d) $x \longmapsto \dfrac{1}{x}$,

   (e) $x \longmapsto -x$, (f) $x \longmapsto x^3$, (g) $x \longmapsto -\dfrac{1}{x}$, (h) $x \longmapsto \sqrt{x}$.
8. 证明: 四群是两个 2 阶循环群的直积.

*9. 证明: 所有非零复数组成的乘法群是单位圆的旋转群和非零实数的乘法群的直积. (提示: 设 $z = r\mathrm{e}^{\mathrm{i}\theta}$.)

10. 证明: 对任意群 G, H, K, 有 $G \times H$ 与 $H \times G$ 同构, $G \times (H \times K)$ 与 $(G \times H) \times K$ 同构.

6.12 自同构 · 共轭元素

定义 *群 G 同它自身的同构称为 G 的自同构. 于是 G 的自同构 α 就是 G 到自身的一一变换 (G 的双射), 并满足*

$$(xy)\alpha = (x\alpha)(y\alpha), \quad \text{对} G \text{中一切} x, y. \tag{16}$$

定理 22 *任意群 G 的全体自同构构成一个群 A.*

证明 (参看定理 6.) 显然, 恒等变换是自同构, 并且任意两个自同构的乘积也是自同构. 最后, 如果 $x \longmapsto x\alpha$ 是一个自同构, 那么由 (16) 式, 有

$$\begin{aligned}(xy)\alpha^{-1} &= [(x\alpha^{-1}\alpha)(y\alpha^{-1}\alpha)]\alpha^{-1}\\ &= \{[(x\alpha^{-1})(y\alpha^{-1})]\alpha\}\alpha^{-1}\\ &= (x\alpha^{-1})(y\alpha^{-1}).\end{aligned}$$

所以 α^{-1} 是一个自同构. 证毕

一个平行的定义和定理可用到整环上, 实际上, 上述内容一般可应用到抽象代数系统上. 我们可以把一个抽象代数系统 A 的自同构恰恰看作 A 的对称.

定义 *在任意群 G 中, $a^{-1}xa$ 称为元素 x 在由 a 共轭变换之下的共轭元素 (或简称 x 在 a 之下的共轭).*

在定理 18 中我们已经看出, 在置换群中, 任意循环的共轭是另一个具有相同长度的循环. 在任意变换群中也有一个类似的说法. 例如, 如果 α 和 ϕ 是空间 S 到自身的一一变换, 则 $\psi = \alpha^{-1}\phi\alpha$ 与 ϕ 的关系如同定理 18 中所说的那样. 特别地, S 中的任意点 q 可以写为 $q = p\alpha$, 这里 p 是 S 中某一点, 并有

$$(p\alpha)\phi = p\alpha(a^{-1}\phi\alpha) = (p\alpha\alpha^{-1})\phi\alpha = (p\phi)\alpha.$$

于是 ψ 是变换 $p\alpha \longmapsto (p\phi)\alpha$; 换句话说, ϕ 在 α 之下的共轭 $\psi = \alpha^{-1}\phi\alpha$ 是由 ϕ 按下面方式得到: 每个点 p 和它的像 $r = p\phi$ 分别用 $p\alpha$ 和 $r\alpha$ 代替. 例如, 在正方形群中,$V = R^{-1}HR$ 表明, 关于垂直轴的反射是关于水平轴的反射在 R 之下的共轭, 这因为 R 把水平轴映射到垂直轴.

定理 23 *对群 G 的任意固定元素 a, 共轭变换 $T_a : x \longmapsto a^{-1}xa$ 是 G 的一个自同构.*

证明 对所有 x, y, 有

$$(a^{-1}xa)(a^{-1}ya) = a^{-1}(xy)a.$$

形为 $x \longmapsto a^{-1}xa$ 的自同构 T_a 称为*内自同构*. 所有其他自同构称为*外自同构*.

可以验证, 正方形对称群有四个不同的内自同构, 有四个外自同构. 另一方面, 三阶循环群除了恒等变换之外没有内自同构, 但是它有外自同构 $x \leftrightarrow x^2$.

定理 24 *任意群 G 的全体内自同构构成 G 的自同构群的一个子群.*

证明 因为 $b^{-1}(a^{-1}xa)b = (ab)^{-1}x(ab)$, 所以内自同构 T_a 和 T_b 的乘积是内自同构 T_{ab}; 类似地, 因为 $(a^{-1})^{-1}(a^{-1}xa)(a^{-1}) = x$, 所以共轭变换 (或自同构)$T_a$ 的逆是 $T_{a^{-1}}$.

定义 (伽罗瓦 (Galois)) *群 G 的子群 S 是 G 中正规子群当且仅当它在 G 的所有内自同构作用之下不变 (也就是说, S 在包含每个元素的同时, 也必包含这个元素的所有共轭元素).*

正规子群有时称为自共轭子群或不变子群.

例如, 正方形的旋转群是正方形对称群的正规子群; 子群 $[I, R^2]$ 也是正规子群. 再有, 阿贝耳群的每个子群都是正规子群, 因为对所有的 x 和 a, 有 $a^{-1}xa = a^{-1}ax = x$. 还有, 平面的平移群是平面所有刚体运动的欧几里得群的一个正规子群 (参看第 9 章).

定理 25 *任意同态 $\theta : G \to H$ 的核 N 是 G 的正规子群.*

证明 根据定理 20 的推论 2 知, N 是 G 的子群. 再有, 如果 $a \in N, b \in G$, 则

$$\theta(b^{-1}ab) = b'^{-1}\theta(a)b' = b'^{-1}e'b' = e',$$

其中 $b' = \theta(b), e' = \theta(e)$, 这是因为, 根据定理 20, 有

$$\theta(b^{-1}) = [\theta(b)]^{-1}.$$

一般地, 设 $a^{-1}Sa$ 表示所有乘积 $a^{-1}sa$(这里 s 在 S 中) 的集合. 那么这个定义表明, S 是正规子群当且仅当对 G 中每个 a, 集合 $a^{-1}Sa$ 等于 S.

定理 26 *子群 S 是正规的当且仅当它的所有右陪集都是它的左陪集.*

证明 如果 S 是正规的, 则对所有的 a 有

$$aSa^{-1} = (a^{-1})^{-1}Sa^{-1} = S;$$

因此 $sa(s \in S)$ 的集合 Sa 同由 $(asa^{-1})a = as(s \in S)$ 组成的集合 $(aSa^{-1})a$ 一样, 于是对所有的 a, 有 $Sa = aS$. 反过来, 如果右陪集 Sa 是左陪集 bS, 则 $a^{-1}Sa = a^{-1}bS$ 包含元素 $e = a^{-1}ea$, 而左陪集 $eS = S$ 也包含元素 e, 根据 6.8 节引理 2, 所以 $a^{-1}Sa = S$.

这个定理的一个推论是, 只有一个陪集的任意子群 S 是正规子群; 不在 S 中的全体元素构成 S 的右陪集和左陪集. 因此交错群是 n 次对称群的正规子群.

注 考虑群 G 的元素 a 和由 a 诱导出的内自同构 T_a 之间的对应. 根据定理 24 的证明, 有 $T_aT_b = T_{ab}$, 这保留了乘法运算. 然而, 如正方形对称群中那样, 这个对应通常不是一一的 (R^2 和 I 诱导出同一个内自同构); 它是一个同态. 我们容易验证, 这个同态的核恰是 G 的中心.

习 题

1. p 阶循环群有多少个自同构? pq 阶循环群呢? 这里 p, q 是不同的素数.
2. 列出四群的全部自同构, 哪些是内自同构?
3. 求出 8 阶循环群的全部自同构.
4. 证明: m 阶循环群的自同构是对应 $a^k \longmapsto a^{rk}$, 这里 r 是环 $\mathbf{Z}_m$ 的单位.
5. 证明: 在任意群中, "x 与 y 共轭" 是一个等价关系.

6. 证明：群的元素 a 诱导出恒等的内自同构当且仅当 a 在群的中心里.

*7. (a) 求出正方形对称群的一个自同构 α, 适合 $R\alpha = R, H\alpha = D$.

(b) 证明 α 是外自同构. (提示：正方形对称群可用 6.6 节中讨论过的生成元 R 和 H 来表示.)

8. 证明：如果 G 和 H 是同构群, 那么 G 和 H 之间不同的同构个数等于 G 的自同构个数.

9. 列举正方形对称群的全体内自同构、共轭元素集合和正规子群.

*10. 设 G 是任意群, A 是 G 的自同构组成的群. 证明：全体偶 (α, g)(其中 $\alpha \in A, g \in G$) 在乘法 $(\alpha, g)(\alpha', g') = (\alpha\alpha', (g\alpha')g')$ 之下构成一个群 (称为 G 的“全形”).

*11. (a) 证明：3 阶循环群的全形是 3 次对称群.

(b) 证明：4 阶循环群的全形是正方形对称群.

12. 证明：如果 M 和 N 都是群 G 的正规子群, 那么它们的交也是 G 的正规子群.

13. 证明：如果 M 和 N 都是群 G 的正规子群, 那么所有乘积 $xy(x \in M, y \in N)$ 构成的集合 MN 是 G 的正规子群.

14. 证明：任意群 G 的全体内自同构是 G 的所有自同构构成的群的正规子群.

*15. (a) 证明：对每个有理数 $c \neq 0$, 对应 $x \longmapsto xc$ 是有理数加法群的自同构.

(b) 证明：有理数加法群没有其他自同构.

*16. 设 G 是 pq 阶 (p, q 为素数) 群. 证明：G 或者是一个循环群, 或者包含一个 p(或 q) 阶元素. 在第二种情形中证明, G 或者包含一个正规子群, 或者包含 q 个 p 阶共轭子群. 对后一种情形, 证明, $pq - q(p-1) = q$ 个非 p 阶元素构成正规子群. 推出 G 总有一个正规真子群.

*17. (a) 证明：如果 $k^n \equiv 1 (\text{mod } m)$, 那么定义关系 $a^m = b^n = e,\ b^{-1}ab = a^k$ 确定了一个具有 m 阶正规子群的 mn 阶群.

(b) 利用习题 16 找出全部 6 阶群和 15 阶群.

*18. 利用习题 16 找出全部可能的 10 阶群和 14 阶群.

*19. 运用习题 16 的分析, 证明：阶数为任意给定素数的平方的群中只有两个不同构.

*6.13 商　　群

现在我们将指出怎样构造某一特定的抽象群 G 的所有同态像 G' 的同构.

诚然, 设 $x \longmapsto x'$ 是群 G 到群 G' 上的任意同态, 并设 N 是这个同态的核. 如果 a 和 b 是 G 的任意元素, 则我们可写成 $b = at$, 因此 $b' = a't'$. 但是根据消去律, $a't' = a'$ 当且仅当 $t' = e'$, 也就是当且仅当 $t \in N$. 总之, $b' = a'$ 当且仅当 $b = at (t \in N)$.

引理 1　*G 的两个元素在 G' 中有同一个像当且仅当它们是在核 N 的同一个陪集 $Nx = xN$ 中.*

这就建立起 G' 的全体元素与核 N 在 G 中的全体陪集之间的一一对应. 因此 G' 的阶等于核 N 在 G 中的陪集的个数 (或称指数).

引理 2　设 x' 和 y' 是 G' 的元素, 那么 $x'y'$ 可按下述方式求出. 设 Nx 和 Ny 分别对应着 x' 和 y', $NxNy$ 是所有乘积 $uv(u \in Nx, v \in Ny)$ 组成的集合, 那么 $x'y'$ 与包含着集合 $NxNy$ 的 N 的 (唯一) 陪集相对应.

证明　如果 $u = ax, v = by(a, b \in N)$, 那么

$$(uv)' = a'x'b'y' = e'x'e'y' = x'y'.$$

于是在同构意义下, G' 可由 G 和 N 来确定, 即 G' 同构于 N 在 G 中的陪集的集合, 这个集合的乘法运算法则是：两个陪集的乘积 $Nx \circ Ny$ 是包含所有乘积 $uv(u \in Nx, v \in Ny)$ 的 (唯一) 陪集.

我们可以用正方形对称群 G 同四群 G'：$[e, a, b, c]$(6.8 节) 之间的同态来说明上述的讨论. 在这个同态之下, $[I, R^2] \longmapsto e, [R, R^3] \longmapsto a, [H, V] \longmapsto b, [D, D'] \longmapsto c$.(从正方形群表可以验证这是一个同态!)$e$ 的原像构成正规子群 $[I, R^2]$, 而其他元素的原像是 $[I, R^2]$ 的陪集. 最后, 通过计算乘积 $[RH, RV, R^3H, R^3V]$, 可以推导出一个典型的运算法则 $ab = c$. 这些乘积在 (实际上是构成)c 的原像的陪集 $[D, D']$ 之中.

反过来, 设 N 是 G 的任意给定的正规子群, 它与任何同态都没有事先的联系. 我们可以由 N 出发构造 G 的同态像 G', 如下所述.

把 G' 的元素定义为 N 的不同的陪集 Nx. N 的任意两个陪集 Nx 和 Ny 的乘积 $Nx \circ Ny$ 定义为包含所有乘积 $uv(u \in Nx, v \in Ny)$ 的集合 $NxNy$ 的陪集. 如果 $u = ax, v = by$, 其中 $a, b \in N$, 则 $uv = axby = ab'xy$, 其中 $b' = xbx^{-1}$ 也在 N 中, 这因为 N 是正规子群. 因此 $N(xy)$ 是一个包含 $NxNy$ 的陪集; 此外, 因为不同的陪集是不相交的, 并且集合 $NxNy$ 是非空的, 因此不存在两个不同的都包含 $NxNy$ 的陪集.

于是我们就对 G' 的全体元素 (又是 G 的所有陪集) 定义了一个单值二元运算, 它可以写成

$$Nx \circ Ny = N(xy). \tag{17}$$

口头上说就是, 任意两个陪集的乘积可以通过在 G 中把任意一对 “代表元素” x 和 y 相乘, 并构成包含乘积 xy 的陪集来求出. 根据公式 (17), 乘积 $Ne \circ Ny = N(ey) = Ny$, 所以陪集 $N = Ne$ 是这个陪集集合的左单位元素, 又因陪集 $(Nx \circ Ny) \circ Nz$ 和 $Nx \circ (Ny \circ Nz)$ 二者都包含 $(xy)z = x(yz)$, 所以陪集的乘法满足结合律. 最后, 陪集 $Nx^{-1} \circ Nx$ 包含元素 $x^{-1}x = e$, 所以必有 $Nx^{-1} \circ Nx = Ne = N$, 因此陪集的左逆元素存在. 这些结果同定理 4 一起可以证明下面的

引理 3 *G 的任意正规子群 N 的全体陪集构成一个乘法群.*

定义 *N 的陪集群称为 G 对 N 的商群 (或因子群), 并用 G/N 表示*[①].

由 (17) 式知, 对应 $x \longmapsto Nx$ 是 G 到 G/N 上的一个同态, 并且这个同态的核是 N.

反之, 我们已经看到 (根据引理 2), 对群 G 到群 G' 上的任意同态, 如果同态核是 N, 那么同态像 G' 与商群 G/N 同构. 我们得出

定理 27 *一个给定的抽象群 G 的同态像是 G 对它的不同正规子群的商群 G/N, N 的陪集的乘法通过公式 (17) 来定义.*

注 从群和正规子群来构造商群类似于从整数环来构造模 n 整数环 (1.9 节和 1.10 节). N 的陪集类似于模 n 的剩余类, 如果把 $x \equiv y(\text{mod}\, N)$ 定义为关系 $xy^{-1} \in N$, 则这个关系与关系 $x \equiv y(\text{mod}\, n)$ 平行. $xy^{-1} \in N$ 等价于断言: x 和 y 在 N 的同一个陪集中 (见 6.8 节习题 6).

习 题

1. 列出所有抽象群, 它们是正方形对称群的同态像.
2. 列出所有抽象群, 它们是正六边形对称群的同态像.
3. 证明: 任意群 G 的中心 Z 是 G 的正规子群, G/Z 与 G 的内自同构群同构.
4. 证明: 在 6.8 节的习题 6 中, 由 $x \equiv y(\text{mod}\, S)$ 可推出对所有的 a 有 $ax \equiv ay(\text{mod}\, S)$ 当且仅当 S 是正规子群.
5. 设 G 是所有形为 $2^k3^m5^n$ 的有理数构成的群, 这里指数 k, m, n 为整数, 而 S 是所有数 2^k 构成的乘法群.

   (a) 描述 S 的全体陪集, (b) 描述 G/S.
6. 设 $G \to G'$ 是一个同态, 证明: G' 的任意子群 S' 的所有原像的集合是 G 的子群 S, 并且, 如果 S' 是正规子群, 那么 S 也是正规子群.

*7. 设 S 是群 G 的一个子群, 而 N 是群 G 的正规子群. 证明: 如果 $S \cap N = e$ 且 $S \bigcup N = G$, 那么 G/N 与 S 同构.

*8. 设 G 是一个群, 形为 $x^{-1}y^{-1}xy$ 的元素称为换位子, 证明: 所有这样的换位子的乘积组成的集合 G 构成 G 的正规子群.

*9. 在习题 8 中, 证明: G/C 是阿贝耳群. 最后证明: 如果 N 是 G 的正规子群, 并且 G/N 是阿贝耳群, 那么 N 包含 C.

*10. 群 G 的两个子群 S 和 T, 如果对某个 $a \in G$ 有 $a^{-1}Sa = T$, 则称它们是共轭的. 证明: G 的任意子群 S 和它的共轭的交是 G 的正规子群.

---

① 如果 G 是阿贝耳群, 这个群中的二元运算用 "+" 表示, 那么每个子群 N 是 G 中的正规子群, 这时商群常常称为差群, 并记作 $G - N$.

*11. (a) 证明：如果 M 和 N 是 G 的正规子群, 并有 $M\bigcap N=e$, 那么对一切 $a\in M, b\in N$, 有 $ab=ba$. (提示：证明 $aba^{-1}b^{-1}\in M\bigcap N$.)

*(b) 证明：在 (a) 中, 如果 $M\bigcup N=G$, 那么 $G=M\times N$.

*12. 设 G 是任意群, S 是 G 的任意子群. 对任意 $a\in G$, 设 T_a 是 S 的全体右陪集 Sx 上的置换 $Sx\longmapsto Sxa$. 证明：

(a) 对应 $a\longmapsto T_a$ 是同态.

(b) 同态核是习题 10 所说的正规子群.

*13. 证明：非正规子群的全体陪集在乘法 (17) 意义下不能构成群.

*6.14 等价关系与同余关系

在定义整数之间的关系 $a\equiv b(\mathrm{mod}\ n)$ 时, 在通过数偶的同余关系 $(a,b)\equiv(a',b')$(这个同余关系的意思是 $ab'=a'b$) 来构造有理数时, 还有其他地方, 我们曾断言过, 任何满足自反律、对称律和传递律的关系可以看作与相等是一类关系. 我们现在系统地讲这个断言的意义.

为方便起见, 把满足自反律、对称律和传递律的关系 R, 即对集合 S 的一切元素 a,b,c, 有性质

$$aRa\quad \text{由}aRb\text{可推出}bRa,\quad \text{由}aRb\text{和}bRc\text{可推出}aRc,$$

称为 S 上的*等价关系*. 如果像在陪集的情形 (6.13 节) 中, 我们把 S 的适当子集当作元素处理, 这样等价关系R 就变成了通常的相等关系. 事实上, 如果 a 是 S 的任意元素, 我们则可用 $R(a)$ 表示与 a 等价的所有元素 b 的集合; $b\in R(a)$ 当且仅当 bRa. 这样一些 R- 子集具有各种简单性质.

引理 1 *由 aRb 可推出 $R(a)=R(b)$, 反之亦真.*

证明 首先假定 aRb, 并设 c 是 $R(a)$ 的任意元素. 那么根据定义有 cRa, 因此再根据传递律得 cRb, 这就意味着 $c\in R(b)$. 反过来, 由对称律得 bRa, 所以由 $c\in R(b)$ 推出 $c\in R(a)$, 这就意味着两个集合 $R(a)$ 和 $R(b)$ 具有相同的元素, 因此 $R(a)=R(b)$.

现在假定 $R(a)=R(b)$. 根据自反律有 bRb, 所以 $b\in R(b)$, 因为 $R(a)=R(b)$, 这意味着 $b\in R(a)$, 于是 aRb, 这就完成了引理的证明.

在 R 是整数之间的模 n 同余关系的特殊情形中, 由整数 a 确定的集合 $R(a)$ 就是包含整数 a 的剩余类. 这里, 引理 1 特别给出断言：$a\equiv b(\mathrm{mod}\ n)$ 当且仅当 a 和 b 属于模 n 的同一个剩余类 (参看 1.10 节). 其他的说明留作习题.

再有, 模 n 全部剩余类把整数的整个集合 $\mathbf{Z}$ 分成不相交的子类, 因此可以说这构成了 $\mathbf{Z}$ 的一个“分划”. 一般地, 类 S 的*分划* π 是把 S 划分成子类 $A,B,C,\cdots,$

使得 S 的每个元素属于一个且仅属于一个子类 (子集合). R- 子集总提供这样的一个分划.

引理 2 两个 R- 子集或者相等, 或者它们没有公共元素, 并且所有 R- 子集的全体构成 S 的一个分划.

证明 如果 $R(a)$ 和 $R(b)$ 包含公共元素 c, 于是 cRa 并且 cRb, 那么根据对称律和传递律有 aRb. 根据引理 1 这意味着 $R(a) = R(b)$. 所以, 如果 $R(a) \neq R(b)$, 这两个类不能重迭. 最后, 集合 S 的每个元素 c 在特定的 R- 子集 $R(c)$ 中, 这是因为, 根据自反律有 cRc, 所以 $c \in R(c)$.

引理 2 的逆命题可直接证得. 如果集合 S 被分划 π 分成不相交的子类 $A, B, C, \cdots$, 那么关系 aRb 可以定义为 a 和 b 属于这个分划的同一个子类中, 这就给出 S 上的一个抽象的等价关系 R. 此外, 通过每个元素 a 依这个关系确定的 R- 子集恰是按分划 π 给出的包含 a 的子类. 这些结论可以概括如下:

定理 28 集合 S 上的每个等价关系 R 确定 S 的一个分划 π, 它把 S 分成不相交的 R- 子类. 反过来, S 的每个分划 π 产生一个等价关系 R. 于是存在一个 S 上全体等价关系 R 和 S 的全体分划 π 之间的一一对应 $R \leftrightarrow \pi$, 使得 S 的元素 a 和 b 属于分划 π 的同一个子类当且仅当 aRb.

在讨论一个可容许的相等关系 (1.11 节) 的必要条件时, 我们还需要某个与二元运算有关的 "代换性质". 利用集合 S 上的等价关系 R 和二元运算 $a \circ b = c$, 这个性质可写成形式

$$\text{由 } aRa' \text{ 和 } bRb' \text{ 可推出 } (a \circ b)R(a' \circ b'). \tag{18}$$

这个条件有着确定的理论含义.

事实上, 设 R 是 S 上的任意等价关系, 又设 π 是相应的分划, 它把 S 分成 R-子集 $A, B, C, \cdots$, 同陪集的情况一样, 我们把 R- 子集看作新系统 $\Sigma = S/R$ 的元素. 同商群 (或模 n 剩余类) 的情况一样, 我们可以由 S 中的二元运算来定义 Σ 中的二元运算,

$$\text{在}\Sigma\text{中}, A \circ B = C\text{当且仅当在}S\text{中}, \text{由}a \in A\text{和}b \in B\text{推出}(a \circ b) \in C. \tag{19}$$

性质 (18) 是说, 如果 a 和 a' 两个元素都在 R- 子集 A 中 (即, aRa'), 并且 b 和 b' 都在 R- 子集 B 中, 那么 $(a \circ b)$ 和 $(a' \circ b')$ 都属于同一个 R- 子集中. 于是, 这个运算得到的 R- 子集 C 便由 A 和 B 唯一确定, 并且在 (19) 式意义下, C 就是 A 和 B 的乘积 $A \circ B$. 换句话说, 代换性质 (18) 等价于断言: 定义 (19) 产生一个 R- 子集 (即 Σ) 上的单值二元运算. 这就证明了

定理 29 已知集合 S 上的一个等价关系 R, 定义在 S 上并具有代换性质 (18) 的任意二元运算产生一个 S 的 R- 子集上的如 (19) 式定义的单值二元运算.

例如, 设 R 是整数集合上的模 n 同余关系, 加法和乘法都具有代换性质 (18), 于是定理 29 产生出 $\mathbf{Z}_n$(1.10 节中定义的) 中剩余类的加法和乘法. 更一般地, 定理 29 可以应用到关系 $a-b\in C$ 上, 这里 C 是任意交换环中的任意理想, 甚至可以推广到其他代数系统, 这个系统的运算不一定是二元的. 一般说, 满足定理 29 条件的关系称为"同余关系". 类似地, 同构、自同构和同态等概念可以应用到一般的代数系统. 例如, 如果 G 和 H 是具有三元运算 (a, b, c) 的代数, 那么 G 到 H 上的同态是具有下面性质的 G 到 H 上的映射 θ, 这个性质是, 对 G 中一切 a,b,c 有

$$(a,b,c)\theta=(a\theta,b\theta,c\theta).$$

习　题

1. 下列关系 R 中, 哪些是等价关系? 如果是, 描述一下 R- 子集.
   (a) G 是群, S 是子群, aRb 意味着 $a^{-1}b\in S$.
   (b) G, S 如 (a) 中所述, aRb 意味着 $ba^{-1}\in S$.
   (c) $\mathbf{Z}$是整数环, aRb 意味着 $a-b$ 是素数.
   (d) $\mathbf{Z}$是整数环, aRb 意味着 $a-b$ 是偶数.
   (e) $\mathbf{Z}$是整数环, aRb 意味着 $a-b$ 是奇数.
2. 设 G 是字母 $x_1,\cdots,x_n$ 的置换群, 设 x_iRx_j 意味着对某个 $\phi\in G$ 有 $x_i\phi=x_j$, R 是等价关系吗?G 是怎样作用在每个 R- 子集上?
3. 设 G 是由全体平面的变换 $(x,y)\longmapsto(x+a,y)$ 组成, 设 $(x,y)R(x',y')$ 意味着对某个 $\phi\in G$ 有 $(x,y)\phi=(x',y')$. 在这种情况下, R- 子集是什么?
4. 设 a 和 b 是实数, 设 aRb 意味着 $a-b$ 是 360 的整数倍.
   (a) R 是等价关系吗?
   (b) 它是加法同余关系吗?
   (c) 它是乘法同余关系吗?
   (d) 看作角度的加法和乘法, 从这里可推出些什么?
5. (a) 设 C 是交换环中任意理想, 证明: 关系 $a-b\in C$ 对于加法和乘法是同余关系.
   (b) 如果 R 是交换环上的任意同余关系, 那么, 当加法和乘法用公式 (19) 定义时, 全体 R- 子集构成另一个交换环.
6. 在习题 1(a) 中, 证明: 代换性质 (18) 的一半对任意子群 S 都成立, 并证明 (18) 的另一半成立当且仅当 S 是正规子群.
7. 设 $\circ: S^2\to S$ 是二元运算, R 是 S 上的等价关系. 证明: 如果由 aRa' 推出 $(a\circ b)R(a'\circ b)$ 和 $(b\circ a)R(b\circ a')$, 那么 (18) 式成立.

第 7 章　向量与向量空间

7.1　平面向量

在物理学中出现一些称为*向量*的物理量，它们不单纯是数量，它们除了有数量大小之外还具有方向．例如，平面上的一个平行移动，它的效果不仅依赖于移动的距离，而且还依赖于移动的方向．为方便起见，我们可以把平行移动表示成具有适当长度和方向的箭头 α(图 7-1)．两个这样的平行移动 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$，表示做完一个平移之后再接着做另一个平移，它们的联合效果就是“总”位移 γ．如果先做平移 $\boldsymbol{\alpha}$ 后做平移 $\boldsymbol{\beta}$，而箭头 $\boldsymbol{\beta}$ 的始端位于箭头 $\boldsymbol{\alpha}$ 的终端，那么总位移 $\gamma=\boldsymbol{\alpha}+\boldsymbol{\beta}$ 就是连接 $\boldsymbol{\alpha}$ 的始端和 $\boldsymbol{\beta}$ 的终端的箭头．这是以 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 为边的平行四边形的对角线，这个求 $\boldsymbol{\alpha}+\boldsymbol{\beta}$ 的法则就是所谓的向量加法的*平行四边形法则*．

一个位移 $\boldsymbol{\alpha}$ 可以放大三倍得到新的位移 $3\boldsymbol{\alpha}$，或者取半得到位移 $\frac{1}{2}\boldsymbol{\alpha}$．我们甚至可以构成它的负倍数，例如 $-2\boldsymbol{\alpha}$，它表示一个大小是 $\boldsymbol{\alpha}$ 的两倍，方向与 $\boldsymbol{\alpha}$ 相反的位移．一般地，$\boldsymbol{\alpha}$ 可以乘上任意实数 c 构成新位移 $c\boldsymbol{\alpha}$，当 c 为正数时，$c\boldsymbol{\alpha}$ 的方向与 $\boldsymbol{\alpha}$ 相同，$c\boldsymbol{\alpha}$ 的大小是 $\boldsymbol{\alpha}$ 的 c 倍，而当 c 是负数时，方向必相反．数 c 称为*标量*(或*纯量*)，乘积 $c\boldsymbol{\alpha}$ 称为*“数乘”积*．

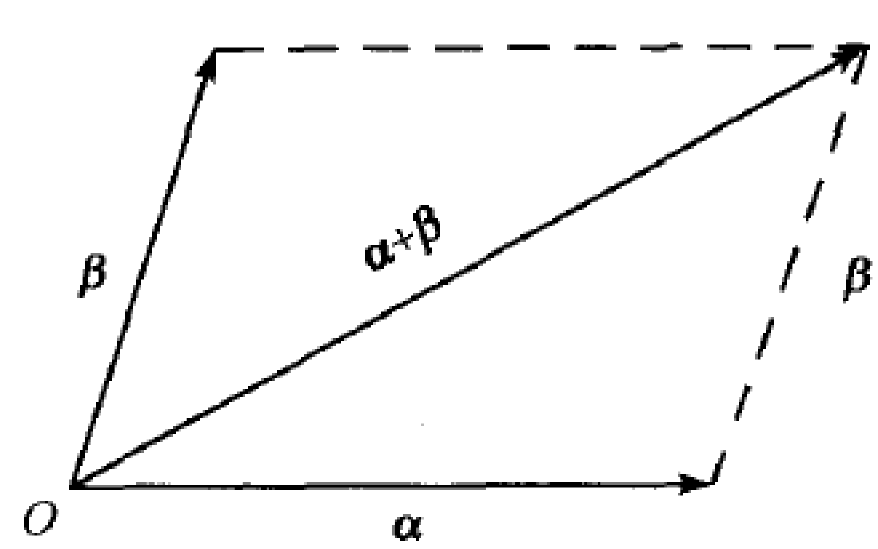


图　7-1

平面中作用于一点的力，以及速度和加速度，都有类似的向量表示，在所有这些情形中，向量加法的平行四边形法则和 (实) 数乘运算，具有同位移情形一样的涵义．这就说明了“各种不同的物理状态可以有相同的数学表示”这样一个原理．

解析几何提出了用实数偶来表示平面向量的方法．我们可以用始端在 $(0,0)$，终端在相应的点 (a_1,a_2) 的箭头 $\boldsymbol{\alpha}$ 表示任何这样的向量，其中坐标 a_1,a_2 是实数．那么向量的和及“数乘”积的坐标可以通过各向量的坐标利用下面法则计算：

$$(a_1,a_2)+(b_1,b_2)=(a_1+b_1,a_2+b_2), \tag{1}$$

$$c(a_1,a_2)=(ca_1,ca_2). \tag{2}$$

从这些法则我们容易得到向量代数[①] 的各种定律，例如

① 本书里我们用小写希腊字母，像 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\cdots,\boldsymbol{\xi},\boldsymbol{\eta},\boldsymbol{\xi},\cdots$ 来表示向量，用小写拉丁字母来表示标量．

$$\alpha + \beta = \beta + \alpha,\ \alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma, \tag{3}$$

$$c(\alpha + \beta) = c\alpha + c\beta, \qquad 1 \cdot \alpha = \alpha, \tag{4}$$

等等. 这些当中有很多 (特别是向量加法的交换律) 还对应着几何原理.

向量运算可以用来表达很多熟悉的几何概念. 例如, 向量 $\alpha = (a_1, a_2)$ 的终端到向量 $\beta = (b_1, b_2)$ 的终端之间的连线中点可以用公式 $\left(\frac{a_1 + b_1}{2}, \frac{a_2 + b_2}{2}\right)$ 给出, 因此也就是用向量和 $\frac{1}{2}(\alpha + \beta)$ 给出. 所得到的向量也可以称为向量 α 和 β 的重心. 向量代数的一组完整公设将在 7.3 节中给出, 我们先描述向量的其他一些例子.

习　题

1. 利用法则 (1) 和 (2) 证明向量代数的定律 (3) 和 (4).
2. 画图说明分配律 (4).
3. 证明：全体平面向量构成加法群.
4. 证明：每个平面向量 α 可以唯一地表示成两个向量之和 $\alpha = \beta + \gamma$, 其中 β 是沿 x 轴方向的向量, γ 是沿 y 轴方向的向量.

7.2 推　　广

上节描述的例子可以在两个方面推广. 第一个方面, **维数** (7.1 节的向量是二维的) 可以是任意的. 首先, 从以下事实我们看出维数可以推广. 按照 7.1 节中处理平面位移和平面力的同样的方法可以处理空间中的位移和力, 唯一的差别是, 对于空间的情形, 向量具有三个分量 (x_1, x_2, x_3), 而平面向量具有两个分量.

其次, 在静力学理论中我们可以看出, 作用在刚体上的力能够分解成六个分量：作用在重心上沿三个互相垂直方向的拉力和绕这些垂直轴的三个旋转力矩. 两个力的合力的分量还可以通过各力的分量来计算, 而数乘运算(用实数去乘) 的含义同上面一样.

更一般地, 对任意正整数 n, 全体 n- 数组 $\alpha = (a_1, \cdots, a_n)$ 构成一个 n 维向量空间, 可以把它看作 n 维几何空间. 例如, 直线是形为 $\alpha + t\beta(\alpha, \beta$ 固定, $\beta \neq 0; t$ 是变量) 的元素的集合; $\alpha_1, \cdots, \alpha_m$ 的重心是 $\frac{1}{m}(\alpha_1 + \cdots + \alpha_m)$, 等等 (这将在 9.13 节中叙述). 为了得到完整的几何理论, 我们只须如在 7.10 节中那样引进距离的概念.

第二方面的推广来源于下面的观察：就涉及的代数性质而论, 向量的分量和标量都不一定是实数, 而可以是任意域上的元素. 实际上, 含有复分量的向量在电路理论和电磁学中常常被用到. 而在第 14 章中, 我们是以研究含有有理标量的向量作为代数数论的基础.

前面两段所描述的推广合并起来叙述就是, 对于任意正整数 n(维数) 和任意标量域 F 推广都成立.

例 向量空间 F^n 是以所有 n- 数组 $\boldsymbol{\alpha} = (a_1, \cdots, a_n)$, $\boldsymbol{\beta} = (b_1, \cdots, b_n), \cdots$(其中分量 a_i, b_i 在 F 中) 作为它的元素. F^n 中的加法和数乘运算定义如下:

$$(a_1, \cdots, a_n) + (b_1, \cdots, b_n) = (a_1 + b_1, \cdots, a_n + b_n), \tag{5}$$

$$c(a_1, \cdots, a_n) = (ca_1, \cdots, ca_n). \tag{6}$$

定理 1 *在向量空间 $V = F^n$ 中, 向量加法和数乘运算具有下列性质:*

在加法运算之下 V 是阿贝耳群; (7)

$c(\boldsymbol{\alpha} + \boldsymbol{\beta}) = c\boldsymbol{\alpha} + c\boldsymbol{\beta}, \qquad (c + c')\boldsymbol{\alpha} = c\boldsymbol{\alpha} + c'\boldsymbol{\alpha}$ (分配律) (8)

$(cc')\boldsymbol{\alpha} = c(c'\boldsymbol{\alpha}), \qquad 1 \cdot \boldsymbol{\alpha} = \boldsymbol{\alpha}.$ (9)

证明 我们首先验证关于群的公设. 向量加法满足结合律, 这是因为对任意像上面那样定义的向量 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$, 和任意向量 $\boldsymbol{\gamma} = (c_1, \cdots, c_n)$, 我们有

$$(\boldsymbol{\alpha} + \boldsymbol{\beta}) + \boldsymbol{\gamma} = (a_1 + b_1 + c_1, \cdots, a_n + b_n + c_n) = \boldsymbol{\alpha} + (\boldsymbol{\beta} + \boldsymbol{\gamma}).$$

上式是根据域中加法的结合律 (6.4 节), 对每个 i, 有 $(a_i + b_i) + c_i = a_i + (b_i + c_i)$. 特殊向量 $\mathbf{0} = (0, \cdots, 0)$ 起着单位元素的作用, 而 $-\boldsymbol{\alpha} = (-\alpha_1, \cdots, -\alpha_n)$ 在 $\boldsymbol{\alpha} + (-\boldsymbol{\alpha}) = (-\boldsymbol{\alpha}) + \boldsymbol{\alpha} = \mathbf{0}$ 的意义下是 α 的逆元素. 注意, $-\boldsymbol{\alpha} = (-1)\boldsymbol{\alpha}$ 还是向量 $\boldsymbol{\alpha}$ 与标量 (-1) 的乘积, 而 $\mathbf{0} = 0\boldsymbol{\alpha}$, 对任意 $\boldsymbol{\alpha}$.

因为对每个 i 有 $a_i + b_i = b_i + \alpha_i$, 所以上述群是可交换的. 同样地, 定义 (5) 和 (6) 将分配律 (8) 的两边化为分量所在域中的分配律.

习 题

1. 设 $\boldsymbol{\alpha} = (1, 1, 0)$, $\boldsymbol{\beta} = \left(-\frac{1}{2}, 0, \frac{2}{3}\right)$, $\boldsymbol{\gamma} = \left(0, \frac{1}{4}, 2\right)$, 计算:

   (a) $\boldsymbol{\alpha} + 2\boldsymbol{\beta} + 3\boldsymbol{\gamma}$, (b) $3(\boldsymbol{\alpha} + \boldsymbol{\beta}) - 2(\boldsymbol{\beta} + \boldsymbol{\gamma})$,

   (c) $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ 的重心是什么? (d) 解方程 $6\boldsymbol{\beta} + 5\boldsymbol{\xi} = \boldsymbol{\alpha}$.

2. 设 $\boldsymbol{\alpha} = (1, \mathrm{i}, 0), \boldsymbol{\beta} = (0, 1 - \mathrm{i}, 2\mathrm{i}), \boldsymbol{\gamma} = (1, 2 - \mathrm{i}, 1)$, 计算:

   (a) $2\boldsymbol{\alpha} - \mathrm{i}\boldsymbol{\beta}$,

   (b) $\mathrm{i}\boldsymbol{\alpha} + (1 + \mathrm{i})\boldsymbol{\beta} - (\mathrm{i} + 3)\boldsymbol{\gamma}$,

   (c) 解方程 $\boldsymbol{\alpha} - \mathrm{i}\boldsymbol{\xi} = \boldsymbol{\beta}$.

3. 设 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 如习题 1 和习题 2 所述, 试把线段 $\overline{\boldsymbol{\alpha}\boldsymbol{\beta}}$ 分为 2:1.

*4. 设 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 如习题 2 所述, 你能把线段 $\overline{\boldsymbol{\alpha}\boldsymbol{\beta}}$ 分为 $1 : 2\mathrm{i}$ 吗? 并加以说明.

5. 设 $\mathbf{Z}_3^n$ 是由 n 维向量组成, 其分量属于模 3 整数域. 试问

   (a) $\mathbf{Z}_3^n$ 中有多少向量?

   (b) 关于 $\mathbf{Z}_3^n$ 中的 $\boldsymbol{\alpha} + \boldsymbol{\alpha} + \boldsymbol{\alpha}$, 你能说些什么?

*6. 你能够定义 $\mathbf{Z}_3^n$ 中任意两点间的"中点"吗? 能定义任意三点 (或四点) 的重心吗?(提示：试验数值例子.)

7.3　向量空间与子空间

我们现在来定义向量空间的一般概念. 向量空间实质上就是一个代数系统, 它的元素在向量加法和数乘运算之下组合在一起, 这里的数 (标量) 是属于一个适当的域 F, 对于这两个运算, 7.2 节中列举的法则都成立.

定义　域 F 上的向量空间 V 是满足下面条件的向量的集合: V 中任意两个向量 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 确定一个 (唯一的) 向量 $\boldsymbol{\alpha}+\boldsymbol{\beta}$ 作为和; 任意向量 $\boldsymbol{\alpha} \in V$ 和任意标量 $c \in F$ 确定一个"数乘"积, 它具有性质 (4) 和 (7)~(9).

(法则 (8) 和 (9) 对于所有向量 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$, 以及所有标量 c 和 c' 者成立.)

定理 1 实质上表明, 对任意正整数 n 和任意域 F, F^n 是向量空间. 还有很多无穷维向量空间, 它们在现代数学分析中起着基本的作用.

例如, 设 S 表示所有实变量 x 的函数 $f(x)$ 的集合, $f(x)$ 在区间 $0 \leqslant x \leqslant 1$ 上单值连续. 两个这样的函数 $f(x)$ 和 $g(x)$ 的和 $h(x)=f(x)+g(x)$ 是 S 中的一个函数, 并且 $f(x)$ 与实常数 c 的"数乘"积 $cf(x)$ 也是一个这样的函数. 这些函数不可能用箭头表示, 但是, 它们的加法和数乘运算具有同我们前述例子同样形式的代数性质, 甚至可以认为, 这个集合 S 中的向量在线段 $0 \leqslant x \leqslant 1$ 的每一点上有一个分量 (即函数值!).

再有, 考虑函数 f, 它的定义域是任意集合 S(比如说, 任意平面区域), 它的取值域是域 F, 这就是说, f 赋给每个 $x \in S$ 一个值 $f(x) \in F$. 如果和 $h=f+g$ 及"数乘"积 $h'=cf$ 是用方程 $h(x)=f(x)+g(x)$ 及 $h'(x)=cf(x)$ 对每个 $x \in S$ 分别定义的, 那么所有这样函数 f 的集合构成 F 上的向量空间.

为了与我们所用的群的加法记号相一致, 我们用$\mathbf{0}$表示这个群的单位元素, 它是满足

$$\boldsymbol{\alpha}+\mathbf{0}=\mathbf{0}+\boldsymbol{\alpha}=\boldsymbol{\alpha},\ \text{对一切}\boldsymbol{\alpha} \tag{10}$$

的唯一的"零"向量. 零向量$\mathbf{0}$ 与零标量 0 不应被混淆. 然而, 它们两个却都是单位元素.

事实上, 对所有的 c 和 $\boldsymbol{\alpha}$, 由 (8) 的两个分配律得到

$$c\boldsymbol{\alpha}+0\boldsymbol{\alpha}=(c+0)\boldsymbol{\alpha}=c\boldsymbol{\alpha}=c\boldsymbol{\alpha}+\mathbf{0},$$
$$c\boldsymbol{\alpha}+c\mathbf{0}=c(\boldsymbol{\alpha}+\mathbf{0})=c\boldsymbol{\alpha}=c\boldsymbol{\alpha}+\mathbf{0}.$$

现在消去两边的 $c\boldsymbol{\alpha}$, 我们得到两个公式

$$0\boldsymbol{\alpha}=\mathbf{0},\ \text{对一切}\ \boldsymbol{\alpha};\qquad c\mathbf{0}=\mathbf{0},\ \text{对一切}c. \tag{11}$$

再有,"数乘" 积 $(-1)\boldsymbol{\alpha}$ 在群中充当任意给定向量 $\boldsymbol{\alpha}$ 的逆元素, 这因为

$$\boldsymbol{\alpha}+(-1)\boldsymbol{\alpha}=1\cdot\boldsymbol{\alpha}+(-1)\boldsymbol{\alpha}=[1+(-1)]\boldsymbol{\alpha}=0\boldsymbol{\alpha}=\mathbf{0}.$$

因此

$$\text{在 (加法) 群中, 任意向量 } \boldsymbol{\alpha} \text{ 的逆向量是} (-1)\boldsymbol{\alpha}. \tag{12}$$

由 (11) 和 (12) 得出, 任意向量的 "幂" 的循环子群是由不同整数 n 和 $\boldsymbol{\alpha}$ 的乘积组成.

在普通的三维向量空间 $\mathbf{R}^3$ 中, 位于一个固定平面上通过原点的所有向量构成一个二维向量空间, 它是整个空间的一部分. 类似地, 位于一个固定直线上通过原点的所有向量的集合 S, 在加法和数乘运算之下是封闭的, 因此这个集合也是 $\mathbf{R}^3$ 的 "子空间".

定义 向量空间 V 的子集合 S, 如果它对于 V 中的向量加法和数乘运算也是一向量空间, 那么 S 称为 V 的子空间.

一个非空子集 S 是子空间当且仅当 S 中任意两个向量之和还在 S 中, 并且 S 的任意向量与标量的乘积还在 S 中. 从定义出发可以很容易地验证这个命题. 很显然, 子空间的定义同以前子域和子群的定义相类似. 从几何上讲, "子空间" 只不过是通过原点 O 的线性子空间 (直线、平面等).

例如, 对任何域 F, 形为 $(0, x_2, 0, x_4)$ 的全体向量构成 F^4 的子空间. 还有, 单独一个零向量 $\mathbf{0}$ 是任意向量空间的子空间.

再有, 次数最高是 7 的多项式集合是所有多项式构成的向量空间的子空间, 这里不管多项式的基域是否是实数域. 类似地, 定义在区间 $0 \leqslant x \leqslant 1$ 上的全体连续函数的集合是定义在同一个定义域上所有函数的线性空间的子空间.

在向量空间 V 中, 给定向量 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$, 所有 $\boldsymbol{\alpha}_i$ 的线性组合

$$c_1\boldsymbol{\alpha}_1+\cdots+c_m\boldsymbol{\alpha}_m \qquad (\text{每个 } c_i \text{ 是标量})$$

的集合是子空间, 这是因为对所有向量 $\boldsymbol{\alpha}_i$ 和所有标量 c_i, c_i' 及 c', 恒等式

$$\begin{aligned}&(c_1\boldsymbol{\alpha}_1+\cdots+c_m\boldsymbol{\alpha}_m)+(c_1'\boldsymbol{\alpha}_1+\cdots+c_m'\boldsymbol{\alpha}_m)\\=&(c_1+c_1')\boldsymbol{\alpha}_1+\cdots+(c_m+c_m')\boldsymbol{\alpha}_m,\end{aligned} \tag{13}$$

$$c'(c_1\boldsymbol{\alpha}_1+\cdots+c_m\boldsymbol{\alpha}_m)=(c'c_1)\boldsymbol{\alpha}_1+\cdots+(c'c_m)\boldsymbol{\alpha}_m, \tag{14}$$

都成立. 这就证明了

定理 2 向量空间 V 中任意一组向量的所有线性组合的集合是 V 的子空间.

这个子空间显然是包含所有给定向量的最小子空间, 因此称它为由给定向量生成(或张成)的子空间. 由单个向量 $\boldsymbol{\alpha}_1 \neq \mathbf{0}$ 张成的子空间是所有 "数乘" 积 $c\boldsymbol{\alpha}_1$ 组成

的集合 S_1; 在几何上, S_1 就是通过原点和 $\boldsymbol{\alpha}_1$ 的直线. 类似地, 由两个非共线的向量 $\boldsymbol{\alpha}_1$ 和 $\boldsymbol{\alpha}_2$ 张成的子空间实际上是一个通过 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 和原点的平面.

定理 3　向量空间 V 的任意两个子空间的交 $S \cap T$ 是 V 的子空间.

证明　两个给定的子空间 S 和 T 的交, 定义为既属于 S 又属于 T 的所有向量的集合 $S \cap T$(参见 6.9 节定理 17, 关于两个子群的交). 如果 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 是两个这样的向量, 则它们的和 $\boldsymbol{\alpha}+\boldsymbol{\beta}$ 一定在 S 中 (因为 S 是包含 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 的子空间), 同样也一定在 T 中, 因此它也在交 $S \cap T$ 中. 类似地, 任意向量 $\boldsymbol{\alpha}$ 的 "数乘" 积 $c\boldsymbol{\alpha}$ 在 $S \cap T$ 中.　证毕

再有, 向量空间 V 的任意两个子空间 S 和 T 确定一个集合 $S+T$, 它是由所有和 $\boldsymbol{\alpha}+\boldsymbol{\beta}$ 组成, 其中 $\boldsymbol{\alpha}$ 属于 S, $\boldsymbol{\beta}$ 属于 T. 根据交换律、结合律和分配律 (3) 和 (4), 集合 $S+T$ 是个子空间, 称为 S 和 T 的*线性和*或*线性张成*. 显然, 它包含 S 和 T, 而其他任意同时包含 S 和 T 的子空间 R 都包含它. 因此线性和的概念类似于两个子群的并 (参见 6.8 节). $S+T$ 的这些性质可以叙述为

$$S \subset S+T, \quad T \subset S+T; \text{由 } S \subset R \text{ 和 } T \subset R, \text{ 推出 } S+T \subset R, \tag{15}$$

这里 $S \subset R$ 的意思是子空间 S 包含在子空间 R 中.

习　　题

1. 证明: 在任何向量空间中, 由 $c\boldsymbol{\alpha} = \mathbf{0}$ 可推出或者 $c = 0$, 或者 $\boldsymbol{\alpha} = \mathbf{0}$.
2. 设 $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ 如 7.2 节习题 1 中所述, 计算

$$7\left[2(\boldsymbol{\alpha}-3\boldsymbol{\beta})+\frac{1}{3}(3\boldsymbol{\beta}-6\boldsymbol{\gamma})\right]-2(\boldsymbol{\alpha}-\boldsymbol{\gamma})+5\boldsymbol{\beta}+2\boldsymbol{\alpha}.$$

3. 设 $\boldsymbol{\alpha}, \boldsymbol{\beta}$ 如 7.2 节习题 2 中所述, 计算 $(1+2\mathrm{i})(2\boldsymbol{\alpha}-3\boldsymbol{\beta})-8\boldsymbol{\alpha}-9\mathrm{i}\boldsymbol{\beta}$.
4. 下列 $\mathbf{Q}^n (n \geqslant 2)$ 的子集合中, 哪些组成子空间 (这里 $\boldsymbol{\xi}$ 表示向量 $(x_1, \cdots, x_n)$)?
   (a) 分量 x_1 为整数的所有 $\boldsymbol{\xi}$;　(b) 分量 $x_2 = 0$ 的所有 $\boldsymbol{\xi}$;
   (c) 或者分量 $x_1 = 0$ 或者分量 $x_2 = 0$ 的所有 $\boldsymbol{\xi}$;
   (d) 满足条件 $3x_1 + 4x_2 = 1$ 的所有 $\boldsymbol{\xi}$;　(e) 满足条件 $7x_1 - x_2 = 0$ 的所有 $\boldsymbol{\xi}$.
5. 下列定义在 $0 \leqslant x \leqslant 1$ 上的实函数 $f(x)$ 的集合中, 哪些是 $0 \leqslant x \leqslant 1$ 上所有实函数向量空间的子空间:
   (a) 所有四次多项式;
   (b) 所有四次或低于四次的多项式 (包括 $f(x) = 0$);
   (c) 满足条件 $2f(0) = f(1)$ 的所有函数;
   (d) 满足条件 $0 + f(1) = f(0) + 1$ 的所有函数;
   (e) 所有正函数;
   (f) 对一切 x 满足 $f(x) = f(1-x)$ 的所有函数.

6. 当 D 取作域 F 时, 3.3 节习题 3 所描述的函数集合中, 哪些构成向量空间?
7. 设 S 是 $\mathbf{Q}^3$ 的子空间, 它是由所有形为 $(0, x_2, x_3)$ 的向量组成, 而 T 是由向量 $(1,2,0)$ 和 $(3,1,2)$ 张成的子空间. 哪些向量在 $S \cap T$ 中? 哪些向量在 $S+T$ 中?
8. 在 $\mathbf{Z}_3^3$ 中, 有多少向量是由 $(1,2,1)$ 和 $(2, 1, 1)$ 张成的? 有多少向量是由 $(1,2,1)$ 和 $(2,1,2)$ 张成的?
9. 证明: 在 $\mathbf{Q}^3$ 中, 平面 $x_3 = 0$ 可以由下面每对向量张成: $(1,0,0)$ 和 $(1,1,0)$; $(2,2,0)$ 和$(4,1,0)$; $(3,2,0)$和$(-3,2,0)$.
10. 证明: 如果 S 是由 $\boldsymbol{\xi}_1$ 和 $\boldsymbol{\xi}_2$ 张成, T 是由 η_1, η_2和η_3 张成, 那么 $S+T$ 是由 $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\eta}_1, \boldsymbol{\eta}_2$和 $\boldsymbol{\eta}_3$ 张成. 推广这个结果.
11. 构造 $\mathbf{Z}_2^2$ 的加法表, 并列出它的子空间.
12. 构造 $\mathbf{Z}_2^3$ 的加法表, 并列出它的子空间.
13. 证明: 一对齐次线性方程 $a_1x_1 + \cdots + a_nx_n = 0$, $b_1x_1 + \cdots + b_nx_n = 0$(其中 a_i, b_i, x_i 全都属于 F) 的所有解 $(x_1, \cdots, x_n)$ 的集合是 F^n 的子空间.

*14. 证明: 向量空间公设 $1 \cdot \boldsymbol{\alpha} = \boldsymbol{\alpha}$ 不能从其他公设推出.(提示: 在平面上构造"伪"数乘积 $c \otimes \boldsymbol{\alpha}$, 它是 $c\boldsymbol{\alpha}$ 在固定直线上的投影.)

*15. 证明: 对于向量加法的交换律公设是多余的.(提示: 用两种方法展开 $(1+1)(\boldsymbol{\alpha}+\boldsymbol{\beta})$.)

7.4 线性无关与维数

向量空间或者子空间的维数这一重要几何概念尚待给出抽象的定义. 它将被描述为张成这个空间 (或子空间) 的向量的最少个数.

例如, 普通空间 $\mathbf{R}^3$ 可以由三个向量 $(1,0,0), (0,1,0)$和$(0,0,1)$ 张成, 它们分别是沿三个坐标轴的单位向量 (长度为 1), 但是 $\mathbf{R}^3$ 不能由两个向量张成 (两个非共线向量张成一个通过原点的平面). 因此 $\mathbf{R}^3$ 的维数是 3.

更一般地, 任意 F^n 由 n 个单位向量

$$\begin{aligned} \boldsymbol{\varepsilon}_1 &= (1, 0, \cdots, 0), \\ \boldsymbol{\varepsilon}_2 &= (0, 1, \cdots, 0), \\ &\vdots \\ \boldsymbol{\varepsilon}_n &= (0, 0, \cdots, 1) \end{aligned} \tag{16}$$

张成. 实际上, F^n 中任意向量是这些单位向量的线性组合, 这因为

$$(x_1, \cdots, x_n) = x_1\boldsymbol{\varepsilon}_1 + \cdots + x_n\boldsymbol{\varepsilon}_n. \tag{17}$$

我们将在定理 5 的推论 2 中证明, F^n 不能由少于 n 个向量张成, 因此有理由称 F^n 为域 F 上的 n 维向量空间.

不仅 $\varepsilon_1,\cdots,\varepsilon_n$ 生成整个空间 F^n, 而且, $x_1\varepsilon_1+\cdots+x_n\varepsilon_n=\mathbf{0}$ 当且仅当 $(x_1,\cdots,x_n)=(0,\cdots,0)$, 即当且仅当 $x_1=\cdots=x_n=0$. 这意味着单位向量在下述意义之下是线性无关的.

定义　向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 线性无关 (在 F 上) 当且仅当对 F 中一切标量 c_i,

$$\text{由 } c_1\boldsymbol{\alpha}_1+\cdots+c_m\boldsymbol{\alpha}_m=\mathbf{0} \text{ 推出 } c_1=\cdots=c_m=0. \tag{18}$$

一组向量如果不是线性无关的, 则称它们线性相关.

线性无关的向量组的任意子集合还是线性无关的, 这是定义的明显的结论. 然而, 下面关于线性组合与线性相关之间的关系更为重要.

定理 4　在空间 V 中非零向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 线性相关当且仅当这些向量中的某个向量是它前面几个向量的线性组合.

证明　在向量 $\boldsymbol{\alpha}_k$ 是它前面几个向量的线性组合 $\boldsymbol{\alpha}_k=c_1\boldsymbol{\alpha}_1+\cdots+c_{k-1}\boldsymbol{\alpha}_{k-1}$ 的情况下, 我们立刻有一个线性关系

$$c_1\boldsymbol{\alpha}_1+\cdots+c_{k-1}\boldsymbol{\alpha}_{k-1}+(-1)\boldsymbol{\alpha}_k=\mathbf{0},$$

其中至少有一个系数 (-1) 不为零. 因此根据 (18), 这些向量线性相关.

反之, 假定向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 线性相关, 于是 $d_1\boldsymbol{\alpha}_1+\cdots+d_m\boldsymbol{\alpha}_m=\mathbf{0}$, 选取最大的下标 k, 使得 $d_k\neq 0$, 然后把 $\boldsymbol{\alpha}_k$ 表示成线性组合

$$\boldsymbol{\alpha}_k=(-d_k^{-1}d_1)\boldsymbol{\alpha}_1+\cdots+(-d_k^{-1}d_{k-1})\boldsymbol{\alpha}_{k-1}.$$

除了 $k=1$ 的情形之外, 上式将 $\boldsymbol{\alpha}_k$ 表为它前面几个向量的线性组合. 在 $d_1\boldsymbol{\alpha}_1=\mathbf{0}$(其中 $d_1\neq 0$) 的情形中, 故有 $\boldsymbol{\alpha}_1=\mathbf{0}$, 这同我们给定的向量没有一个为零向量的假定矛盾.　　证毕

例如, 三个向量 $\boldsymbol{\beta}_1=(2,0,0),\boldsymbol{\beta}_2=(1,3,0)$ 和 $\boldsymbol{\beta}_3=(0,-2,0)$ 不能张成整个空间 $\mathbf{R}^3$, 因为它们位于同一个平面上. 我们可以用关系 $\boldsymbol{\beta}_1-2\boldsymbol{\beta}_2-3\boldsymbol{\beta}_3=\mathbf{0}$ 或者 (解出 $\boldsymbol{\beta}_1$) 用关系 $\boldsymbol{\beta}_1=2\boldsymbol{\beta}_2+3\boldsymbol{\beta}_3$ 来表示这个线性相关. 于是集合 $(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)$ 同它的一个真子集 $(\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)$ 张成同一个子空间. 这就证明了

推论 1　一组向量线性相关当且仅当它包含一个真 (即最小) 子集与原向量组张成同一个子空间.

这也就是说, 我们可以从这组向量中, 删去任意一个向量, 它是 $\mathbf{0}$ 或者它是它前面向量的线性组合, 并可证明剩下来的向量生成的子空间与原来一组向量生成的子空间相同. 现在用归纳法, 我们得到

推论 2　任意有限的向量集合包含一个线性无关的子集合, 它张成的子空间与原集合张成的子空间相同.

现在我们可以叙述线性相关的基本定理.

定理 5 设 n 个向量张成向量空间 V, 它包含 r 个线性无关向量, 那么 $n \geqslant r$.

证明 设 $A_0 = [\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n]$ 是张成 V 的 n 个向量的序列, 并设 $X = [\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_r]$ 是 V 的 r 个线性无关向量的序列. 因为 A_0 张成 V, 所以 $\boldsymbol{\xi}_1$ 是 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 的线性组合, 因此序列 $B_1 = [\boldsymbol{\xi}_1, \boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n]$ 张成 V, 而且是线性相关的. 根据定理 4, B_1 的某一个向量必与它前面的向量线性相关. 这个向量不可能是 $\boldsymbol{\xi}_1$, 因为 $\boldsymbol{\xi}_1$ 属于线性无关向量组 X, 因此在 B_1 中, 某个向量 $\boldsymbol{\alpha}_i$ 依赖于它前面的向量 $\boldsymbol{\xi}_1, \boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_{i-1}$. 像推论 1 那样, 删去这个向量 $\boldsymbol{\alpha}_i$, 子序列 $A_1 = [\boldsymbol{\xi}_1, \boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_{i-1}, \boldsymbol{\alpha}_{i+1}, \cdots, \boldsymbol{\alpha}_n]$ 仍然张成 V.

现在重复论证. 构造序列 $B_2 = [\boldsymbol{\xi}_2, A_1] = [\boldsymbol{\xi}_2, \boldsymbol{\xi}_1, \boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_{i-1}, \boldsymbol{\alpha}_{i+1}, \cdots, \boldsymbol{\alpha}_n]$, 同 B_1 一样, B_2 张成 V, 而且它是线性相关的. 因此和前面一样, B_2 中某一个向量是它前面向量的线性组合. 因为这些 ξ_i 是线性无关的, 所以这个向量不会是 $\boldsymbol{\xi}_2$ 或 $\boldsymbol{\xi}_1$, 故一定是某个 $\boldsymbol{\alpha}_j$, 其中下标 $j \neq i$(比如说, $j > i$). 删去这个 $\boldsymbol{\alpha}_j$, 剩下可张成 V 的 n 个向量的新序列

$$A_2 = [\boldsymbol{\xi}_2, \boldsymbol{\xi}_1, \boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_{i-1}, \boldsymbol{\alpha}_{i+1}, \cdots, \boldsymbol{\alpha}_{j-1}, \boldsymbol{\alpha}_{j+1} \cdots, \boldsymbol{\alpha}_n]$$

这一论证可以重复 r 次, 直到 X 的元素都取完. 每一次失去 A_0 的一个元素. 因此 A_0 最初就一定至少包含 r 个元素, 这就证明了 $n \geqslant r$. 证毕

定理 5 有几个重要推论. 虽然它们包含的"基"和"维数"等概念的完整涵义直到 7.8 节才变得显然, 然而为方便起见, 我们现在还是证明这些推论.

定义 生成 (张成) 整个向量空间的一组线性无关向量称为这个向量空间的基. 一个向量空间是有限维的当且仅当它有一组有限基.

例如, (16) 式的单位向量 $\varepsilon_1, \cdots, \varepsilon_n$ 是 F^n 的一组基.

推论 1 任意有限维向量空间 V 的一切基都包含着相同数目 (有限个) 的元素.

证明 因为 V 是有限维的, 所以它有一组有限基

$$A = [\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n],$$

设 B 是 V 的任意另一组基, 因为 A 张成 V, 并且 B 是线性无关的, 定理 5 表明 B 是有限的, 比如说有 r 个元素, 于是 $n \geqslant r$. 另一方面, B 张成 V, 而 A 是线性无关的, 因此 $r \geqslant n$, 所以 $n = r$.

有限维向量空间 V 的任意一组基中元素的个数称为 V 的维数, 并用 $d[V]$ 表示. 根据定理 5 我们有

推论 2 如果向量空间 V 的维数是 n, 那么, (i) V 的任意 $n+1$ 个元素是线性相关的; (ii) $n-1$ 个元素的任意集合不可能张成 V.

定理 6 有限维向量空间 V 的任意一组线性无关向量是 V 的基的一部分.

证明 设这个线性无关向量组是 $\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_r$, 并设 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 是 V 的一组基. 构成序列 $C=[\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_r,\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n]$. 我们从 C 中可以抽出一个线性无关的子序列 (定理 4 的推论 2), 它也张成空间 V(因而它是 V 的一组基), 这个子序列是通过删去那些是它前面向量的线性组合的向量而得到的. 因为这些 $\boldsymbol{\xi}_i$ 是线性无关的, 所以没有一个 ξ_i 被删去, 因此所得到的这组基包含每一个 $\boldsymbol{\xi}_i$.

推论 n 维向量空间 V 的 n 个向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 是一组基的充分条件是: 或者它们张成空间 V, 或者它们线性无关.

证明 如果 $A=[\boldsymbol{\alpha}_1\cdots,\boldsymbol{\alpha}_n]$ 张成 V, 则它包含一个子集合 A', 这个 A' 是 V 的一组基 (定理 4 的推论 2); 因为 V 的维数是 n, 所以 A' 必有 n 个元素 (定理 5 的推论 1), 因此 $A'=A$, 于是 A 是 V 的一组基. 再有, 如果 A 是线性无关的, 那么根据定理 6, A 是基的一部分, 根据定理 5 的推论 1, 这组基应有 n 个元素, 所以 A 本身就一定是一组基.

习 题

1. 证明: 在 F^2 中, 向量 (a_1,a_2) 和 (b_1,b_2) 线性相关当且仅当 $a_1b_2-a_2b_1=0$
2. 向量 $(1,1,0)$ 和 $(0,1,1)$ 构成 $\mathbf{Q}^3$ 的一组基吗? 为什么?
3. 证明: 如果 $\boldsymbol{\beta}$ 不在子空间 S 中, 而在由 S 和 $\boldsymbol{\alpha}$ 张成的子空间中, 那么 $\boldsymbol{\alpha}$ 在由 S 和 $\boldsymbol{\beta}$ 张成的子空间中.
4. 证明: 如果 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3$ 在 $\mathbf{R}^n$ 中线性无关, 那么 $\boldsymbol{\xi}_1+\boldsymbol{\xi}_2,\boldsymbol{\xi}_1+\boldsymbol{\xi}_3,\boldsymbol{\xi}_2+\boldsymbol{\xi}_3$ 也线性无关. 这个结论在每个 F^n 中都正确吗?
5. 由 $\mathbf{Z}_3^7$ 中四个线性无关的元素张成的每个子空间中有多少个元素? 推广你的结论.
6. 定义整环 D 上的向量空间, 在这个更一般的情形中, 迄今所讨论的公设和定理中有哪些不能成立?

*7. 证明: 具有有理坐标的三个向量在 $\mathbf{Q}^3$ 中线性无关当且仅当它们在 $\mathbf{R}^3$ 中线性无关. 从两个方面推广这个结果.

8. 设向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 线性无关, 证明: 向量 $\boldsymbol{\beta}$ 是 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 的线性组合当且仅当向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}$ 线性相关.

*9. 证明: 实数 $1,\sqrt{2}$ 和 $\sqrt{5}$ 在有理数域上线性无关.

10. 在 $\mathbf{C}^3$ 中找出四个向量, 它们一起张成二维子空间, 并且它们之中任意两个向量线性无关.
11. 证明: 如果 $c_1\boldsymbol{\alpha}+c_2\boldsymbol{\beta}+c_3\boldsymbol{\gamma}=\mathbf{0}$, 其中 $c_1c_3\neq 0$, 那么 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 生成的子空间与 $\boldsymbol{\beta}$ 和 $\boldsymbol{\gamma}$ 生成的子空间相同.
12. 证明: 如果向量空间 V 的两个子空间 S 和 T 具有相同的维数, 那么由 $S\subset T$ 可推出 $S=T$.

*13. (a)$\mathbf{Z}_2^3$ 中有多少两元素的线性无关组? 有多少三元素的线性无关组? 有多少四元素的

线性无关组?

(b) 把你的公式推广到 $\mathbf{Z}_2^n$ 上和 $\mathbf{Z}_p^n$ 上.

*14. $\mathbf{Z}_p^n$ 有多少不同的 k 维子空间.

7.5 矩阵与行等价

与 F^n 中含有数值坐标的向量集合有关的问题, 差不多总可以描述为联立线性方程组问题. 这样, 它们常常可以用 2.3 节中叙述的消去法求解. 我们现在就开始系统地研究这个方法, 这个方法是以矩阵及其行等价等基本概念为中心的. 我们首先给出矩阵的定义.

定义 *在域 F 上, m 行和 n 列元素组成的长方阵列称为 F 上的 $m\times n$ 矩阵.*

注 显然, 任意域 F 上的全体 $m\times n$ 矩阵在下述两种运算之下构 mn 维向量空间: (i) 用同一个标量 c 去乘矩阵的所有元素; (ii) 两矩阵各对应分量相加.

我们现在运用矩阵的概念来确定, 在什么情况下 F^n 的两组向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 和 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$ 张成同一个子空间. 显然, 向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 确定一个 $m\times n$ 矩阵

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix}, \tag{19}$$

它的第 i 行由向量 $\boldsymbol{\alpha}_i$ 的 n 个分量 $a_{i1},\cdots,a_{in}$ 组成. 矩阵 (19) 可以缩写为 (a_{ij}). 矩阵 $\boldsymbol{A}$ 的每一行看作 F^n 的向量, 称为行向量, 由全体行向量张成的 F^n 的子空间称为矩阵 $\boldsymbol{A}$ 的*行空间*. 我们现在要问: 什么时候两个 $m\times n$ 矩阵有相同的行空间呢? 也就是说, 什么时候它们的行向量在 F^n 中张成相同的子空间呢? 这个问题的部分答案是通过我们现在要定义的行等价的概念给出的.

我们现在考虑下列称为*初等行运算*的三个典型步骤作用在 (19) 式矩阵 $\boldsymbol{A}$ 时的效果:

(i) 任意两行互换;

(ii) 某一行元素乘以 F 中任意非零常数 c;

(iii) 某一行的任意倍数加到其他任意一行上.

如果 $m\times n$ 矩阵 $\boldsymbol{B}$ 可以从 $m\times n$ 矩阵 $\boldsymbol{A}$ 通过有限次初等行运算得到, 那么称 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ *行等价*. 因为每一个这样的运算的效果可以通过另一个同类型运算抵消, 使矩阵不变, 因此我们有下面引理.

引理 *任何初等行运算的逆仍是初等行运算.*

因此, 如果 $\boldsymbol{B}$ 行等价于 $\boldsymbol{A}$, 那么 $\boldsymbol{A}$ 行等价于 $\boldsymbol{B}$, 也就是说, 行等价关系是对称的. 显然具有自反性和传递性, 因此它是一个等价关系.

定理 7　*行等价矩阵具有相同的行空间.*

证明　用 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 表示 $m\times n$ 矩阵 $\boldsymbol{A}$ 的逐个行向量. 那么 $\boldsymbol{A}$ 的行空间是所有形为 $c_1\boldsymbol{\alpha}_1+\cdots+c_m\boldsymbol{\alpha}_m$ 的向量组成的集合, 并且初等行运算变为:

(i) $\boldsymbol{\alpha}_i$ 与 $\boldsymbol{\alpha}_j$ 互换 $(i\neq j)$;

(ii) 对任意标量 $c\neq 0$, 用 $c\boldsymbol{\alpha}_i$ 代替 $\boldsymbol{\alpha}_i$;

(iii) 对任意 $j\neq i$ 和任意标量 d, 用 $\boldsymbol{\alpha}_i+d\boldsymbol{\alpha}_j$ 代替 $\boldsymbol{\alpha}_i$.

只须考虑每种类型单个初等行运算作用在行空间上的效果. 因为类型 (i) 和 (ii) 的运算显然不改变行空间, 所以我们只注意类型 (iii) 的单个初等行运算的情形. 取一个典型的情况, 即把第二行的倍数加到第一行上, 这就是把 $\boldsymbol{A}$ 的各行向量分别用行等价矩阵 $\boldsymbol{B}$ 的各新行向量

$$\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1+d\boldsymbol{\alpha}_2,\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\beta}_m=\boldsymbol{\alpha}_m \tag{20}$$

来代替. $\boldsymbol{B}$ 的行空间的任意向量 γ 具有形式 $\gamma=\Sigma c_i\boldsymbol{\beta}_i$, 因此把 (20) 代入, 我们有

$$\gamma=c_1(\boldsymbol{\alpha}_1+d\boldsymbol{\alpha}_2)+c_2\boldsymbol{\alpha}_2+\cdots+c_m\boldsymbol{\alpha}_m,$$

这表明 γ 在 $\boldsymbol{A}$ 的行空间中. 反过来, 根据引理, A 的行向量可以通过 $\boldsymbol{B}$ 的行向量表示为

$$\boldsymbol{\alpha}_1=\boldsymbol{\beta}_1-d\boldsymbol{\beta}_2,\boldsymbol{\alpha}_2=\boldsymbol{\beta}_2,\cdots,\boldsymbol{\alpha}_m=\boldsymbol{\beta}_m,$$

所以同样的论证指出 $\boldsymbol{A}$ 的行空间包含在 $\boldsymbol{B}$ 的行空间中, 于是两个行空间相等.

上述证明立即得到

推论 1　*把矩阵 $\boldsymbol{A}$ 化为行等价矩阵 $\boldsymbol{B}$ 的任意一系列初等行运算可以明显地把 $\boldsymbol{B}$ 的行向量表示为 $\boldsymbol{A}$ 的行向量的线性组合.*

联立线性方程组　下面我们应用行等价矩阵的概念重新说明一下 2.3 节所描述的"高斯消去法". 考虑联立线性方程组

$$\begin{aligned}
a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n&=a_{1,n+1},\\
a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n&=a_{2,n+1},\\
\vdots\qquad\vdots\qquad\qquad\vdots\quad&\quad\vdots\\
a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n&=a_{m,n+1},
\end{aligned} \tag{21}$$

这里系数 a_{ij} 是域 F 中已知常数, 我们想要知道什么样的解向量 $\boldsymbol{\xi}=(x_1,x_2,\cdots,x_n)$ (如果存在的话) 满足已知方程组 (21).

容易验证, 满足 (21) 的一组解向量 $\boldsymbol{\xi}$ 在下列各种运算之下是不变的:

(i) 任意两个方程互换.

(ii) 一个方程乘上 F 中任意非零常数 c.

(iii) 一个方程的任意倍数加到其他任意一个方程上.

但是这些运算应用到 (21) 中的 $m\times(n+1)$ 常数矩阵 (a_{ij}) 时, 它们恰是前面定义过的三种初等行运算. 这就证明了

推论 2 如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是同一个域 F 上的 $m\times(n+1)$ 行等价矩阵, 那么联立线性方程组 (21) 同方程组

$$\begin{aligned} b_{11}x_1+b_{12}x_2+\cdots+b_{1n}x_n&=b_{1,n+1},\\ b_{21}x_1+b_{22}x_2+\cdots+b_{2n}x_n&=b_{2,n+1},\\ \vdots\qquad\vdots\qquad\quad&\quad\vdots\qquad\vdots\\ b_{m1}x_1+b_{m2}x_2+\cdots+b_{mn}x_n&=b_{m,n+1}. \end{aligned}\tag{21$'$}$$

有相同解向量 $\boldsymbol{\xi}=(x_1,\cdots,x_n)$ 集合.

习 题

1. 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是行等价矩阵, 证明: $\boldsymbol{A}$ 的行向量线性无关当且仅当 $\boldsymbol{B}$ 的行向量线性无关.
2. 证明: 如果运算 (iii) 用下面 (iii$'$) 来代替, 那么行等价的涵义不变.
   (iii$'$) 任意一行加到其他任意一行.

*3. 证明: 任意 (i) 类初等行运算可以用四次 (ii)、(iii) 类运算来完成. (提示: 用 2×2 矩阵试验.)

7.6 线性相关的检验

现在我们的目的在于使用初等行运算把已知的 $m\times n$ 矩阵 $\boldsymbol{A}$ 尽可能简化. 在 $\boldsymbol{A}$ 的任意非零的行中, 第一个非零元素称为这一行的 "首" 元素. 如果矩阵 $\boldsymbol{A}$ 满足下列两个条件, 则我们称 $\boldsymbol{A}$ 是*行简化矩阵*:

(a) 每个首元素 (非零行的) 是 1.

(b) 包含这样的首元素的每列中, 其他元素都是零.

4×6 行简化矩阵的例子是

$$\begin{pmatrix} 0&0&1&r_{14}&r_{15}&r_{16}\\ 1&0&0&r_{24}&r_{25}&r_{26}\\ 0&1&0&r_{34}&r_{35}&r_{36}\\ 0&0&0&0&0&0 \end{pmatrix},\qquad \begin{pmatrix} 1&d_{12}&0&d_{14}&0&d_{16}\\ 0&0&1&d_{24}&0&d_{26}\\ 0&0&0&0&1&d_{36}\\ 0&0&0&0&0&0 \end{pmatrix}.\tag{22}$$

定理 8 任意矩阵 $\boldsymbol{A}$, 可通过 (ii) 类和 (iii) 类初等行运算使它行等价于行简化矩阵.

证明 假设含有元素 a_{ij} 的已知矩阵 $\boldsymbol{A}$, 它的第一行非零, 其首元素是 a_{1t}, 位于第 t 列. 用 a_{1t}^{-1} 乘第一行, 这行的首元素变为 1. 现在对每个 $i \neq 1$, 从第 i 行减去第一行的 a_{it} 倍. 这就把第 t 列的其他每个元素化为零, 因此对于第一行, 条件 (a) 和 (b) 满足.

现在按照同样的方式逐次处理其他各行. 在处理第 k 行时, 含有第 $1, \cdots, k-1$ 各行首元素的列中的元素不变, 这是因为这些列和第 k 行交叉的元素都已变为零. 因此当处理完第 k 行之后, 我们得到的矩阵其前 k 行满足条件 (a) 和 (b). 对 k 用归纳法就推出定理 8.

通过行的置换 (即相继进行 (i) 类初等行运算), 显然我们可以重新排列行简化矩阵 $\boldsymbol{R}$ 的各行, 使得

(c) $\boldsymbol{R}$ 的每个零行都排在 $\boldsymbol{R}$ 的所有非零行的下面.

假定有 r 个非零行, 对于 $i = 1, 2, \cdots, r,$, 第 i 行首元素出现在 t_i 列. 因为所有这样的列中其他元素都为零, 所以当 $i \neq j$ 时我们有 $t_i \neq t_j$. 再通过行的置换, 我们可重新排列 $\boldsymbol{R}$ 使得

(d) $t_1 < t_2 < \cdots < t_r$(第 i 行首元素在第 t_i 列中).

如果行简化矩阵还满足 (c) 和 (d), 则称为 (行)*简化梯形矩阵*(首元素位于“梯”上). 我们已证明了:

推论 任意矩阵同简化梯形矩阵行等价.

例如, (22) 式的第二个矩阵已经是简化梯形矩阵; (22) 式的第一个矩阵却不是, 但是, 把第一行放在第三行下面就可化成简化梯形矩阵.

定理 9 设 $\boldsymbol{E}$ 是含有非零行 $\gamma_1, \cdots, \gamma_r$ 的行简化矩阵, 各行首元素 1 位于第 $t_1, \cdots, t_r$ 列. 那么对 $\boldsymbol{E}$ 的行空间中的任意向量 $\boldsymbol{\beta}$ 有

$$\boldsymbol{\beta} = y_1\gamma_1 + \cdots + y_r\gamma_r,$$

其中 γ_i 的系数 y_i 是 $\boldsymbol{\beta}$ 的第 t_i 列的元素, 也就是 $\boldsymbol{\beta}$ 的第 t_i 个分量.

证明 因为 $\boldsymbol{E}$ 的第 t_i 列元素除了 γ_i 行那个元素是 1 外, 其余都是 0, 所以 $\boldsymbol{\beta}$ 的第 t_i 个分量一定是 $y_i \cdot 1$.

推论 1 行简化矩阵的全体非零行向量线性无关.

这是因为如果 $\boldsymbol{\beta} = \boldsymbol{0}$, 则根据上述定理每个 $y_i = 0$.

推论 2 设 $m \times n$ 矩阵 $\boldsymbol{A}$ 与行简化矩阵 $\boldsymbol{R}$ 行等价, 那么 $\boldsymbol{R}$ 的全体非零行向量构成 $\boldsymbol{A}$ 的行空间的一组基.

证明 根据推论 1, $\boldsymbol{R}$ 的这些行向量线性无关, 并张成 $\boldsymbol{R}$ 的行空间. 于是它们是这个行空间的一组基, 根据定理 7, $\boldsymbol{R}$ 的行空间与 $\boldsymbol{A}$ 的行空间恒等. 证毕

矩阵 $\boldsymbol{A}$ 的行空间的维数称为矩阵 $\boldsymbol{A}$ 的**秩**, 记作 $\operatorname{rank}(\boldsymbol{A})$. 因为这个空间是由 $\boldsymbol{A}$ 的全体行向量张成, 这些行向量一定包含张成行空间的一组线性无关的行向量, 所以我们看到, $\boldsymbol{A}$ 的秩也可被描述成 $\boldsymbol{A}$ 的线性无关行向量的最大数目. 根据定理 7, 行等价矩阵具有相同的秩.

特别是, $n\times n$ 矩阵 (方阵)$\boldsymbol{A}$ 的秩为 n 当且仅当它的所有行向量线性无关. 主对角线 (从左上方到右下方) 上的元素都为 1, 而其他元素都为 0 的 $n\times n$ 矩阵称为 $n\times n$**单位矩阵**, 记作 $\boldsymbol{I}_n$.

推论 3 *一个 $n\times n$ 矩阵的秩为 n 当且仅当它与 $n\times n$ 单位矩阵 $\boldsymbol{I}_n$ 行等价.*

证明 设 $\boldsymbol{A}$ 与秩为 n 的简化梯形矩阵 $\boldsymbol{E}$ 行等价, 则矩阵 $\boldsymbol{E}$ 有 n 个非零行向量, 因此 n 个首元素 1 在 n 个不同的列中, 在这些列中除了首元素外没有其他非零元素 (这些列包括所有的列). 适当调整行序, 则 $\boldsymbol{E}$ 恰好是单位矩阵. 证毕

在检验向量线性无关时, 或更一般地, 在计算子空间的维数 (等于矩阵的秩) 时, 不一定使用简化梯形矩阵, 只须把矩阵化为任意梯形矩阵就可以了, 例如下面的 4×7 矩阵的形式

$$\boldsymbol{E}=\begin{pmatrix} 0 & 1 & d_{13} & d_{14} & d_{15} & d_{16} & d_{17} \\ 0 & 0 & 0 & 1 & d_{25} & d_{26} & d_{27} \\ 0 & 0 & 0 & 0 & 1 & d_{36} & d_{37} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

于是梯形矩阵可以通过下面条件来定义: 每一非零行的首元素是 1, 第一行后的每一行中, 首元素 1 前零的个数大于前面那些行首元素前零的个数.

这样, 化成梯形矩阵之后, 利用下面的定理可直接求出矩阵的秩.

定理 10 *任意矩阵 $\boldsymbol{A}$ 的秩是任意行等价于 $\boldsymbol{A}$ 的梯形矩阵的非零行的个数.*

证明将留作习题.

例 检验 $\boldsymbol{\alpha}_1=(1,-1,1,3)$, $\boldsymbol{\alpha}_2=(2,-5,3,10)$ 和 $\boldsymbol{\alpha}_3=(3,3,1,1)$ 的线性无关性.

通过 (iii) 类初等行运算得到新的行向量 $\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1$, $\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-2\boldsymbol{\alpha}_1=(0,-3,1,4)$, $\boldsymbol{\beta}_3=\boldsymbol{\alpha}_3-3\boldsymbol{\alpha}_1=(0,6,-2,-8)$. 最后, 设 $\boldsymbol{\gamma}_1=\boldsymbol{\beta}_1$, $\boldsymbol{\gamma}_2=-\frac{1}{3}\boldsymbol{\beta}_2$, $\boldsymbol{\gamma}_3=\boldsymbol{\beta}_3-6\boldsymbol{\gamma}_2=\boldsymbol{\beta}_3+2\boldsymbol{\beta}_2=\mathbf{0}$. 结果得到含有行向量 $\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3$ 的梯形矩阵 $\boldsymbol{C}$,

$$\boldsymbol{C}=\begin{pmatrix} 1 & -1 & 1 & 3 \\ 0 & 1 & -\dfrac{1}{3} & -\dfrac{4}{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

因为 $\boldsymbol{C}$ 有零行, 所以原来的向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性相关. 把前面的关系代入 $\boldsymbol{\gamma}_3=\mathbf{0}$ 中, 我们得出 α_i 之间明显的依赖关系

$$\mathbf{0}=\boldsymbol{\gamma}_3=\boldsymbol{\beta}_3+2\boldsymbol{\beta}_2=(\boldsymbol{\alpha}_3-3\boldsymbol{\alpha}_1)+2(\boldsymbol{\alpha}_2-2\boldsymbol{\alpha}_1)=-7\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3.$$

行等价的附录　简化梯形矩阵为行等价性的检验提供了方便的方法.

定理 11　只存在一个 $m\times n$ 简化梯形矩阵 $\boldsymbol{E}$ 具有已知行空间 $S\subset F^n$.

证明　设具有行空间 S 的简化梯形矩阵 $\boldsymbol{E}$ 有非零行向量 $\gamma_1,\cdots,\gamma_r$, 这里 γ_i 的首元素是 1, 位于第 t_i 列中. 由条件 (d), 有 $t_1<t_2<\cdots<t_r$. 设 $\boldsymbol{\beta}=y_1\gamma_1+\cdots+y_r\gamma_r$ 是 $\boldsymbol{E}$ 的行空间中任意非零向量; 根据定理 9, 向量 β 的第 t_i 个分量为 y_i, 如果 y_s 是 $y_1,\cdots,y_r$ 中第一个非零元素, 那么 $\boldsymbol{\beta}=y_s\gamma_s+\cdots+y_r\gamma_r$. 因为 $t_s<\cdots<t_r$, 所以剩下行向量 $\gamma_{s+1},\cdots,\gamma_r$ 的首元素在第 t_s 列后面的那些列中, 因此 $\boldsymbol{\beta}$ 有 y_s 作为它的首元素, 位于第 t_s 列中. 换句话说, S 中的每个向量 $\boldsymbol{\beta}$ 具有首元素位于第 $t_1,\cdots,t_r$ 列中的一列. 这些列的每一列都出现 (γ_i 的首元素所在的列). 因此行空间 S 确定了指标 $t_1,\cdots,t_r$.

$\boldsymbol{E}$ 的行向量 $\gamma_1,\cdots,\gamma_r$ 中, 每一行都有首元素 1, 在第 $t_1,\cdots,t_r$ 列中 (对某一行而言), 除一列外其余各列都是零元素. 如果 $\boldsymbol{\beta}$ 是 S 的任意向量, 它的首元素 1 在某一列 (第 t_i 列) 中, 其他各列 (第 t_j 列) 的元素为零, 那么根据定理 9, $\boldsymbol{\beta}$ 一定是 γ_i. 于是行空间和这些列指标唯一确定了 $\boldsymbol{E}$ 的行向量 $\gamma_1,\cdots,\gamma_r$. 满足定理要求.

证毕

推论 1　任意 $m\times n$ 矩阵 $\boldsymbol{A}$ 与一个且仅与一个简化梯形矩阵行等价.

这个结果容易证明. 它还可以概括为如下说法: 简化梯形矩阵给出行等价意义下的矩阵标准型. 也就是说, 每个矩阵与一个且仅与一个特殊的标准型矩阵行等价.

推论 2　两个 $m\times n$ 矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 行等价当且仅当它们有同一个行空间.

证明　如果 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 行等价, 那么根据定理 7, $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 有同一个行空间. 反之, 如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 有同一个行空间, 它们分别行等价于简化梯形矩阵 $\boldsymbol{E}$ 和 $\boldsymbol{E}'$. 因为 $\boldsymbol{E}$ 和 $\boldsymbol{E}'$ 有同一个行空间, 根据定理 11, $\boldsymbol{E}$ 和 $\boldsymbol{E}'$ 相等. 因此 (通过 $\boldsymbol{E}=\boldsymbol{E}'$), $\boldsymbol{A}$ 确实与 $\boldsymbol{B}$ 行等价.

这些结果再次表明, 矩阵的行等价恰是研究 F^n 子空间的另一种语言.

习　题

1. 证明:

$$\begin{pmatrix}5&2&7\\-3&4&1\\-1&-2&-3\end{pmatrix} 与 \begin{pmatrix}1&0&1\\0&1&1\\0&0&0\end{pmatrix} 行等价.$$

2. 把下列各矩阵化为行等价的梯形矩阵:

(a) $\begin{pmatrix}1&-1&3\\2&-4&1\\0&3&2\end{pmatrix}$,　　(b) $\begin{pmatrix}-5&6&-3\\3&1&11\\4&-2&8\end{pmatrix}$;

(c) $\begin{pmatrix} 1 & 6 & -2 & 5 \\ 4 & 0 & 4 & -2 \\ 7 & 2 & 0 & 2 \\ -6 & 3 & -3 & 3 \end{pmatrix}$, (d) $\begin{pmatrix} 2 & -1 & 3 & 2 \\ 0 & 2 & 1 & 4 \\ 4 & -2 & 3 & 9 \\ 2 & -3 & 4 & 5 \end{pmatrix}$,

(e) $\begin{pmatrix} \mathrm{i} & 1 & -\mathrm{i} & 1+\mathrm{i} \\ 1 & -\mathrm{i} & \mathrm{i} & 2-\mathrm{i} \\ -1 & 0 & 1 & 0 \\ 2 & \mathrm{i} & 2\mathrm{i} & 3\mathrm{i} \end{pmatrix}$.

3. 在习题 2 中, 把梯形矩阵的各行向量表示成原来矩阵各行向量的线性组合.
4. 检验下列各组向量是否线性相关:

   (a) (1, 0, 1), (0, 2, 2), (3, 7, 1) 在 $\mathbf{Q}^3$ 中或在 $\mathbf{C}^3$ 中.

   (b) (0,0,0), (1, 0, 0), (0, 1, 1) 在 $\mathbf{R}^3$ 中.

   (c) (1, i, 1+i), (i, −1, 2 − i), (0, 0, 3) 在 $\mathbf{C}^3$ 中.

   (d) (1, 1, 0), (1, 0, 1) (0, 1, 1) 在 $\mathbf{Z}_2^3$ 中和在 $\mathbf{Z}_3^3$ 中.

   在线性相关的各种情况中, 取出生成相同行空间的线性无关子集.
5. 在 $\mathbf{Q}^6$ 中检验下列各组向量的线性无关性, 并找出张成子空间的基:

   (a) $(2,4,3,-1,-2,1),(1,1,2,1,3,1),(0,-1,0,3,6,2)$.

   (b) $(2,1,3,-1,4,-1),(-1,1,-2,2,-3,3),(1,5,0,4,-1,7)$.
6. 把习题 5 中两组向量放在一起, 找出张成子空间的基.
7. 求出下列矩阵的秩和矩阵行空间的基:

   (a) $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}$, (b) $\begin{pmatrix} 1 & 2 & 1 & 2 \\ 3 & 2 & 3 & 2 \\ -1 & -3 & 0 & 4 \\ 0 & 4 & -1 & -3 \end{pmatrix}$, (c) $\begin{pmatrix} 1 & 2 & 4 & 5 & 7 \\ 1 & 2 & 3 & 4 & 5 \\ -1 & -2 & 0 & 2 & 1 \end{pmatrix}$.
8. 列出含有两个非零行向量的 2×4 简化梯形矩阵的所有可能形式.(这些产生"格拉斯曼 (Grassmann) 流形"的胞腔部分, 这里流形的点是四维空间中通过原点的平面.)
9. 证明: $m \times n$ 矩阵的秩, 既不超过 m 也不超过 n.
10. 如果 $m \times (n+k)$ 矩阵 $\boldsymbol{B}$ 是由 $m \times n$ 矩阵 $\boldsymbol{A}$ 再添上 k 个新的列构成的, 那么 rank $(\boldsymbol{A}) \leqslant \operatorname{rank}(\boldsymbol{B})$.
11. 直接证明 (不用定理 8): 任意矩阵 $\boldsymbol{A}$ 行等价于一个梯形矩阵 (不一定是简化梯形矩阵).

7.7 向量方程 · 齐次方程

当我们想要求解形为

$$\lambda = x_1\boldsymbol{\alpha}_1 + \cdots + x_m\boldsymbol{\alpha}_m \quad (\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m\text{为}F^n\text{中固定向量}, \lambda\text{为任意向量}) \qquad (23)$$

的一些向量方程时, 用矩阵代替线性方程组 (21), 并对矩阵使用初等行运算, 这样做

是特别方便的.

例如, 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 是 7.6 节的例中给出的向量, 设 $\boldsymbol{\lambda}=(2,7,-1,-6)$. 把矩阵 $\boldsymbol{A}$ 化为梯形矩阵 $\boldsymbol{C}$, 我们先解方程

$$\boldsymbol{\lambda}=y_1\boldsymbol{\gamma}_1+y_2\boldsymbol{\gamma}_2+y_3\boldsymbol{\gamma}_3=y_1\boldsymbol{\gamma}_1+y_2\boldsymbol{\gamma}_2.$$

比较等式两边的第一个分量, 我们得 $y_1=2$; 比较第二个分量, 我们则得 $7=-y_1+y_2$ 即 $y_2=9$. 因此, 如果 λ 确实是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 的线性组合, 那么我们一定有

$$\boldsymbol{\lambda}=2\boldsymbol{\gamma}_1+9\boldsymbol{\gamma}_2=2\boldsymbol{\alpha}_1-3\boldsymbol{\beta}_2=2\boldsymbol{\alpha}_1-3(\boldsymbol{\alpha}_2-2\boldsymbol{\alpha}_1)=8\boldsymbol{\alpha}_1-3\boldsymbol{\alpha}_2,$$

计算 $8\boldsymbol{\alpha}_1-3\boldsymbol{\alpha}_2$ 的第三分量和四分量, 我们看出 $\boldsymbol{\lambda}$ 的确是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 的线性组合.

因为 $\boldsymbol{\gamma}_3=-7\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3=\mathbf{0}$, 所以在上述情况下, (23) 的另一些解是

$$\boldsymbol{\lambda}=(8-7y)\boldsymbol{\alpha}_1+(-3+2y)\boldsymbol{\alpha}_2+y\boldsymbol{\alpha}_3,$$

其中 y 是任意的. 这确实是 (23) 的最一般的解. 如果向量 $\boldsymbol{\lambda}$ 换成 $\boldsymbol{\lambda}'=(2,7,1,-6)$, 则上面过程指出 $\boldsymbol{\lambda}'$ 根本不可能表示为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 的线性组合.

事实上; 当含有几个向量 $\boldsymbol{\lambda}$ 时, 常常最好是把含有行向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 的 $m\times n$ 矩阵变换成含有非零行向量 $\gamma_1,\cdots,\gamma_r$ 的简化梯形矩阵 $\boldsymbol{C}$. 因为矩阵的每个初等行运算只包含有限次有理运算 (即加、减、乘、除), 又因为经过有限次初等行运算之后可以把给定的矩阵变换成简化梯形矩阵, 所以经有限次有理运算之后, 可以把给定的矩阵变换成简化梯形矩阵.

那么, 应用定理 9 我们可以得到一组唯一可能的系数 $y_1,\cdots,y_r$ 使得 $\boldsymbol{\lambda}=y_1\gamma_1+\cdots+y_r\gamma_r$. 如果这个方程不是对 γ 的所有分量都成立, 那么 $\boldsymbol{\lambda}$ 就不在 $\boldsymbol{A}$ 的行空间中, 因此 (23) 就没有解. 如果这个方程对 $\gamma_1,\cdots,\gamma_r$ 的所有分量都成立, 那么, 因为 $\boldsymbol{C}$ 的各行向量都是 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 的线性组合 $\boldsymbol{\gamma}_i=\sum\limits_{j=1}^{m}e_{ij}\boldsymbol{\alpha}_j$, 所以我们得到 (23) 的解为 $\boldsymbol{\lambda}=\sum y_ie_{ij}\boldsymbol{\alpha}_j$, 因此我们有 $x_j=y_1e_{1j}+\cdots+y_re_{rj}$. 这就证明了下面的结果.

定理 12 对 F^n 中已知向量 $\boldsymbol{\lambda},\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$, 向量方程 $\boldsymbol{\lambda}=x_1\boldsymbol{\alpha}_1+\cdots+x_m\boldsymbol{\alpha}_m$ 可以通过 F 中的有限次有理运算解出 (如果解存在的话).

推论 设 S 和 T 分别是由向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 和 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_k$ 张成的 F^n 的子空间, 那么关系式 $S\supset T,T\supset S$ 和 $S=T$ 可以通过有限次有理运算来检验.

这是因为, 我们可以由 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 经过初等行运算来构造一组非零向量 $\boldsymbol{\gamma}_1,\cdots,\boldsymbol{\gamma}_r$, 它们就是简化梯形矩阵的行向量, 而且还张成 S. 然后我们像上面那样检验 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_k$ 是否都是 $\gamma_1,\cdots,\gamma_r$ 的线性组合, 显然这是 $S\supset T$ 的充分必要条件. 把前

面的过程反过来, 我们可以确定是否 $T \supset S$. 这两个过程结合起来可以检验 $S = T$ 是否成立. 另外还可以把以 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 为行向量的矩阵和以 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k$ 为行向量的矩阵都变换成简化梯形矩阵, 来检验 $S = T$ 是否成立, 因为 $S = T$ 成立当且仅当它们的简化梯形矩阵具有相同的非零向量.

简化梯形矩阵还可以用来确定形为

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0, \\ \vdots \qquad \vdots \qquad\qquad \vdots \quad & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned} \tag{24}$$

的齐次线性方程组的解. 我们设 S 是 F^n 中所有满足 (24) 的向量 $\boldsymbol{\xi} = (x_1, \cdots, x_n)$ 的集合. 容易验证 S 是一个子空间. 我们将指出如何确定这个子空间的基.

首先看到, 同 2.3 节一样, 用初等行运算作用在方程组 (24) 上可以把它变换成等价的方程组, 特别是, 当作用到 $m \times n$ 矩阵 $\boldsymbol{A}$(它的第 i 行是 (24) 的第 i 个方程的系数 $(a_{i1}, \cdots, a_{in})$) 时, 这些运算把 $\boldsymbol{A}$ 变成具有相同 "解向量" $\boldsymbol{\xi} = (x_1, \cdots, x_n)$ 的集合 S 的另一个矩阵. 现在把 $\boldsymbol{A}$ 化为简化梯形矩阵, 其中首元素都为 1, 位于第 $t_1, \cdots, t_r$ 列上. 相应的方程组有 r 个非零方程, 并且第 i 个方程是含有未知数 x_{t_i} 的唯一的方程.

为使记号简单, 假定首元素出现在前 r 列 (事实上, 只要对未知数 $x_1, \cdots, x_n$, 也就是对 $\boldsymbol{A}$ 的各列, 作适当的置换, 这是可以办到的). 那么化简后的方程组具有形式

$$\begin{aligned} x_1 + c_{1,r+1}x_{r+1} + \cdots + c_{1n}x_n &= 0, \\ x_2 + c_{2,r+1}x_{r+1} + \cdots + c_{2n}x_n &= 0, \\ \vdots \qquad \vdots \qquad\qquad \vdots \quad & \\ x_r + c_{r,r+1}x_{r+1} + \cdots + c_{rn}x_n &= 0. \end{aligned} \tag{25}$$

在这个简化了的形式中, 任意选取 $x_{r+1}, \cdots, x_n$ 的值, 并对 $x_1, \cdots, x_r$ 解方程组 (25). 得到解向量

$$\boldsymbol{\xi} = \left(-\sum_{j=r+1}^{n} c_{1j}x_j, \cdots, -\sum_{j=r+1}^{n} c_{rj}x_j, x_{r+1}, \cdots, x_n \right), \tag{26}$$

显然我们就可以得到方程组 (25) 的一切解. 特别是, 在参数 $x_{r+1}, \cdots, x_n$ 中, 令其中一个为 1, 其他为 0, 我们就得到 $n - r$ 组解

$$\begin{aligned} \boldsymbol{\xi}_{r+1} &= (-c_{1,r+1}, \cdots, -c_{r,r+1}, 1, 0, \cdots, 0), \\ &\ \ \vdots \qquad\qquad \vdots \qquad\quad \vdots \\ \boldsymbol{\xi}_n &= (-c_{1n}, \cdots, -c_{rn}, 0, 0, \cdots, 1). \end{aligned}$$

这 $n-r$ 个解向量是线性无关的 (因为前 r 个坐标全都忽略, 它们是线性无关的). 公式 (26) 表明, 一般解 $\boldsymbol{\xi}$ 刚好是这 $n-r$ 个基本解的线性组合 $\boldsymbol{\xi}=x_{r+1}\boldsymbol{\xi}_{r+1}+\cdots+x_n\boldsymbol{\xi}_n$. 于是我们就找到了已知方程组 (24) 的解向量空间 S 的一组基, 因此证明了

定理 13 含有 n 个未知数的 r 个线性无关的齐次线性方程组, 它的所有解 $(x_1,\cdots,x_n)$ 组成的"解空间"的维数为 $n-r$.

推论 含有 n 个未知数 $x_1,\cdots,x_n$ 的 n 个线性无关的齐次线性方程组的唯一解是

$$x_1=x_2=\cdots=x_n=0.$$

例 设 S 是由方程 $x_1+x_2=x_3+x_4$ 和 $x_1+x_3=2(x_2+x_4)$ 定义的. 这样, 在几何上, S 是四维空间中两个三维超平面的交. 这些方程的矩阵化简如下

$$\begin{pmatrix}1&1&-1&-1\\1&-2&1&-2\end{pmatrix}\to\begin{pmatrix}1&1&-1&-1\\0&-3&2&-1\end{pmatrix}\to\begin{pmatrix}1&4&-3&0\\0&-3&2&-1\end{pmatrix}.$$

最后一个矩阵 (除了符号和列序外) 是简化梯形矩阵. 这就得出等价方程组 $x_1+4x_2-3x_3=0,\ -3x_2+2x_3-x_4=0$, 它具有一般解为

$$\boldsymbol{\xi}=(3x_3-4x_2,x_2,x_3,-3x_2+2x_3),$$

令 $x_2=0,x_3=1$ 和 $x=1,x_3=0$ 得到解空间的一组基 (3, 0, 1, 2) 和 $(-4,1,0,-3)$.

根据对偶原则, 我们可以得到由任意子空间的一切向量所满足的线性方程组的基. 例如, 设 T 是 F^4 中由向量 $(1,1,-1,-1)$ 和 $(1,-2,1,-2)$ 张成的子空间. 那么齐次线性方程 $\Sigma a_ix_i=0$ 对 T 中所有向量 (x_1,x_2,x_3,x_4) 是恒等式当且仅当 $a_1+a_2=a_3+a_4$ 和 $a_1+a_3=2(a_2+a_4)$. 满足这些方程的系数向量 (a_1,a_2,a_3,a_4) 的集合的基前面已经求出, 把那里的 x 用 a 代替.

上述例子的线性方程 $x_1+x_2-x_3-x_4=0$ 和 $x_1-2x_2+x_3-2x_4=0$ 等价于向量方程

$$x_1(1,1)+x_2(1,-2)+x_3(-1,1)+x_4(-1,-2)=(0,0).$$

它的解是二维空间 F^2 中四个向量 $(1,1),(1,-2),(-1,1),(-1,-2)$ 之间所有线性依赖关系. 它还可以像 7.5 节那样, 把以这四个向量作为行向量的 4×2 矩阵化简成梯形矩阵来求解, 这个矩阵可以从前面的 2×4 矩阵经过转置行和列而得到.

习 题

1. 设 $\boldsymbol{\xi}_1=(1,1,1)$, $\boldsymbol{\xi}_2=(2,1,2)$, $\boldsymbol{\xi}_3=(3,4,-1)$, $\boldsymbol{\xi}_4=(4,6,7)$, 求出不全为零的数值 c_i 满足 $c_1\boldsymbol{\xi}_1+c_2\boldsymbol{\xi}_2+c_3\boldsymbol{\xi}_3+c_4\boldsymbol{\xi}_4=\mathbf{0}$.

2. 设 $\eta_1=(1+i,2i)$, $\eta_2=(2,-3i)$, $\eta_3=(2i,3+4i)$, 求出所有满足 $c_1\eta_1+c_2\eta_2+c_3\eta_3=\mathbf{0}$ 的复数 c_i.
3. 求出两个向量, 使它们张成由所有满足 $x_1+x_2=x_3-x_4=0$ 的向量 (x_1,x_2,x_3,x_4) 组成的子空间.
4. 求出两个向量, 使它们张成由所有满足

$$3x_1-2x_2+4x_3+x_4=x_1+x_2-3x_3-2x_4=0$$

的向量 (x_1,x_2,x_3,x_4) 组成的子空间.
5. 求出下列各方程组解向量空间的基:

(a) $x+y+3z=0,$
$2x+2y+6z=0;$

(b) $x+y+z=0,$
$y+z+t=0;$

(c) $x+2y-4z=0,$
$3x+y-2z=0;$

(d) $x+y+z+t=0,$
$2x+3y-z+t=0,$
$3x+4y+2t=0.$

6. 把习题 5 中的方程式换成模 5 同余式, 求同余式组解向量空间的基.
7. 确定下列各向量方程 (在有理数域上) 是否有解. 如果有解, 就求出它的一组解.
   (a) $(1,-2)=x_1(1,1)+x_2(2,3)$,
   (b) $(1,1,1)=x_1(1,-1,2)+x_2(2,1,3)+x_3(1,-1,0)$,
   (c) $(2,-1,1)=x_1(2,0,3)+x_2(3,1,2)+x_3(1,2,-1)$.
8. 在 $\mathbf{Q}^4$ 中, 设 $\alpha_1=(1,1,2,2)$, $\alpha_2=(1,2,3,4)$, $\alpha_3=(0,1,3,2)$ 和 $\alpha_4=(-1,1,-1,1)$. 把下列各向量表示成形式 $x_1\alpha_1+x_2\alpha_2+x_3\alpha_3+x_4\alpha_4$:

(a) $(1,0,1,0)$, (b) $(3,-2,1,-1)$, (c) $(0,1,0,0)$, (d) $(2,-2,2,-2)$,

9. 证明: 对 $m\times n$ 矩阵至多进行 m^2 次初等行运算就可以把它变换成行简化矩阵.
10. 证明: 对 4×6 矩阵至多进行 56 次乘法、42 次加减法和 4 次互换运算 (像 $aa^{-1}=1$, $a-a=0$ 或 $0\cdot a=0$ 都没有计算在内) 就可变换成行简化矩阵.

*11. 对 $n\times n$ 矩阵叙述并证明类似于习题 10 的结果.

7.8 基与坐标系

我们已经把张成向量空间 V 的一组线性无关向量定义为空间 V 的基. 基的实际意义在于, 空间 F^n 的任意基的向量在适当选取的坐标系下可以看作空间的单位向量. 这个证明依赖于下面的定理.

定理 14 如果 $\alpha_1,\cdots,\alpha_n$ 是 V 的一组基, 那么 V 的每个向量 ξ 可唯一地表示成 $\alpha_1,\cdots,\alpha_n$ 的线性组合

$$\xi=x_1\alpha_1+\cdots+x_n\alpha_n. \tag{27}$$

证明 因为 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 是 V 的一组基, 它们张成 V, 所以 V 中的每个向量至少有一种方式表示成 (27) 的形式. 如果某个向量 $\boldsymbol{\xi}\in V$ 有第二种这样的表示式 $\boldsymbol{\xi}=x_1'\boldsymbol{\alpha}_1+\cdots+x_n'\boldsymbol{\alpha}_n$, 那么从 (27) 中减去它, 重新合并一下就得出

$$\mathbf{0}=\boldsymbol{\xi}-\boldsymbol{\xi}=(x_1-x_1')\boldsymbol{\alpha}_1+\cdots+(x_n-x_n')\boldsymbol{\alpha}_n.$$

因为 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 是一组基, 它们是线性无关的, 所以从上面的等式推出 $x_1-x_1'=\cdots=x_n-x_n'=0$, 因此每个 $x_i=x_i'$, 于是表示式 (27) 是唯一的.

我们把 (27) 式中的标量 x_i 称为向量 ξ 关于基 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 的坐标. 如果

$$\boldsymbol{\eta}=y_1\boldsymbol{\alpha}_1+\cdots+y_n\boldsymbol{\alpha}_n$$

是 V 中第二个向量, 其坐标为 $y_1,\cdots,y_n$, 那么由向量代数的恒等式, 有

$$\boldsymbol{\xi}+\boldsymbol{\eta}=(x_1+y_1)\boldsymbol{\alpha}_1+\cdots+(x_n+y_n)\boldsymbol{\alpha}_n. \tag{28}$$

口头上说就是, 向量和关于任意基的坐标可以通过把被加向量相应的坐标相加来求得. 类似地, 形为 (27) 的向量 $\boldsymbol{\xi}$ 与标量 c 的乘积是

$$c\boldsymbol{\xi}=c(x_1\boldsymbol{\alpha}_1+\cdots+x_n\boldsymbol{\alpha}_n)=(cx_1)\boldsymbol{\alpha}_1+\cdots+(cx_n)\boldsymbol{\alpha}_n, \tag{29}$$

所以 $c\boldsymbol{\xi}$ 的每个坐标是 c 和 $\boldsymbol{\xi}$ 的相应坐标的乘积.

类似于整环同构和群同构定义, 现在我们定义同一域 F 上的两个向量空间 V 和 W 之间的同构 C: $V\to W$ 是, 由 V 到 W 上适合下面条件的一一对应 $\boldsymbol{\xi}\longmapsto\boldsymbol{\xi}C$:

$$\begin{aligned}&(\boldsymbol{\xi}+\boldsymbol{\eta})C=\boldsymbol{\xi}C+\boldsymbol{\eta}C \text{ 和 } (c\boldsymbol{\xi})C=c(\boldsymbol{\xi}C),\\ &\qquad(\text{对}V\text{中一切向量}\boldsymbol{\xi},\boldsymbol{\eta},\text{对}F\text{中一切标量}c),\end{aligned} \tag{30}$$

那么公式 (28) 和 (29) 表明, F 上的向量空间 V 的每一组基 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 提供了一个由 V 到 F^n 上的同构. 这个同构就是对应 C_α, 它赋给 V 中每个向量 $\boldsymbol{\xi}$ 关于基 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 的坐标的 n- 数组, 即

$$(x_1\boldsymbol{\alpha}_1+\cdots+x_n\boldsymbol{\alpha}_n)C_\alpha=(x_1,\cdots,x_n)\in F^n. \tag{31}$$

因为基向量的个数 n 是由空间的维数 n 确定, 而它是不变的 (定理 5 的推论 1), 所以我们就证明了

定理 15 *域 F 上的任意有限维向量空间与一个且只与一个空间 F^n 同构.*

这样我们就解决了确定 (精确到同构) 所有有限维向量空间的问题. 而且我们已经指出, 同一向量空间的一切基在同构意义下是等价的, 即存在 V 的一个同构, 它把任意一组基映射到其他任意一组基.

一个向量空间可以有很多不同的基. 例如, 根据定理 7, 我们从 $\varepsilon_1, \cdots, \varepsilon_n$ 出发逐次使用初等行运算而得到 F^n 的任何一组向量, 它是 F^n 的一组基. 特别地, 对任何使 $1+1 \neq 0$ 的域 F, $\boldsymbol{\alpha}_1 = (1,1,0)$, $\boldsymbol{\alpha}_2 = (0,1,1)$, 和 $\boldsymbol{\alpha}_3 = (1,0,1)$ 是 F^3 的一组基. 同样, 在普通三维空间中任意三个非共面向量确定了"斜角坐标"向量的一组基.

再有, 如果在复数域 $\mathbf{C}$ 中只考虑复数加法和复数的数乘 (用实数乘) 而不管其他所有代数运算, 那么复数域 $\mathbf{C}$ 可以看作实数域 $\mathbf{R}$ 上的向量空间. 这个空间的维数是 2, 因为 1 和 i 构成一组基, 它们分别生成实数和纯虚数的子空间. $1+\mathrm{i}$ 和 $1-\mathrm{i}$ 两个数构成 $\mathbf{R}$ 上空间 $\mathbf{C}$ 的另一组基, 但这组基用起来不方便.

另外, 考虑齐次线性微分方程

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} - 3\frac{\mathrm{d}x}{\mathrm{d}t} + 2x = 0.$$

不难验证, 这个方程的两个解的和 $x_1(t) + x_2(t)$ 还是方程的解, 一个解同任意 (实) 常数的乘积也是方程的解. 因此这个微分方程的所有解组成的集合 V 是一个向量空间, 有时称它为微分方程的"解空间". 描述这个空间的最容易的办法是说, e^t 和 e^{2t} 构成这个解空间的一组基, 这就意味着一般解可以唯一地表示成形式 $x = c_1\mathrm{e}^t + c_2\mathrm{e}^{2t}$.

最后, 域 F 上关于未定元 x 的所有多项式形式构成的整环 $F[x]$ 是 F 上的向量空间, 因为在 $F[x]$ 中向量空间的一切公设都满足. 多项式相等的定义用于方程 $p(x) = 0$, 就意味着所有的幂 $1, x, x^2, x^3, \cdots$ 在 F 上线性无关. 因此这些幂组成了 $F[x]$ 的一组无穷基, 因为任意向量 (多项式形式) 可以表示成这组基的有限子集的线性组合.

在 $\mathbf{R}^3$ 中, 通过原点的平面 S 和通过原点但不在 S 中的直线 T 张成整个空间, 所以空间中任意向量可以唯一地表示成这个平面上的一个向量与这条直线上的一个向量的和. 更一般地, 设 S 和 T 是向量空间 V 的子空间, 如果 V 的每个向量 ξ 可以唯一地表示成 S 的一个向量和 T 的一个向量的和:

$$\boldsymbol{\xi} = \boldsymbol{\sigma} + \boldsymbol{\tau}, \quad \boldsymbol{\sigma} \in S, \quad \boldsymbol{\tau} \in T \tag{32}$$

那么我们称 V 是两个子空间S 和 T 的*直和*(直接和).

因为 $(\boldsymbol{\sigma}+\boldsymbol{\tau}) + (\boldsymbol{\sigma}'+\boldsymbol{\tau}') = (\boldsymbol{\sigma}+\boldsymbol{\sigma}') + (\boldsymbol{\tau}+\boldsymbol{\tau}')$, 所以对应 $(\boldsymbol{\sigma}, \boldsymbol{\tau}) \longmapsto (\boldsymbol{\sigma}+\boldsymbol{\tau})$ 是由向量空间 V 的加法群映上到 S 和 T 的加法群的*直积* (6.11 节) 的一个同构. 更一般地, F^n 是 F 的加法群的 n 重直积 (作为加法群), 记作 $F^n = F \times F \times \cdots \times F$($n$ 个因子).

反过来, 如果 S 和 T 是同一个域 F 上的任意给定的两个向量空间, 那么我们可以定义一个新的向量空间 $V = S \oplus T$, 它的加法群是 S 和 T 的加法群的直积,

它的数乘运算由公式 $c(\eta,\zeta)=(c\eta,c\zeta)$(对一切 $c\in F$) 来定义. 在这个空间 V 中, $(\eta,0)$ 和 $(0,\zeta)$ 组成的子集合分别构成与 S 和 T 同构的子空间, 而且按上面的定义, V 是这两个子空间的直和. 我们也把 $S\oplus T$ 说成已知向量空间 S 和 T 的直和.

定理 16　*如果有限维向量空间 V 是它的子空间 S 和 T 的直和, 那么 S 的任意基和 T 的任意基的并就是 V 的一组基.*

证明　设 S 和 T 的基分别是 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_k$ 和 $\gamma_1,\cdots,\gamma_m$; 我们希望证明 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_k,\gamma_1,\cdots,\gamma_m$ 是 V 的一组基. 首先, 这些向量张成 V, 这因为 V 中任意向量 $\boldsymbol{\xi}$ 可写成 $\boldsymbol{\xi}=\boldsymbol{\eta}+\boldsymbol{\zeta}$, 其中 $\boldsymbol{\eta}$ 是 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_k$ 的线性组合, $\boldsymbol{\zeta}$ 是 $\gamma_1,\cdots,\gamma_m$ 的线性组合. 其次, 这些向量是线性无关的, 这因为如果

$$\mathbf{0}=b_1\boldsymbol{\beta}_1+\cdots+b_k\boldsymbol{\beta}_k+c_1\gamma_1+\cdots+c_m\gamma_m, \tag{33}$$

那么 $\mathbf{0}$ 就表示成 S 中向量 $\boldsymbol{\eta}_0=\Sigma b_i\boldsymbol{\beta}_i$ 与 T 中向量 $\boldsymbol{\xi}_0=\Sigma c_j\gamma_j$ 之和. 但是 $\mathbf{0}=\mathbf{0}+\mathbf{0}$ 是 $\mathbf{0}$ 作为 S 的向量与 T 的向量之和的另一表示. 根据假设, 表示是唯一的, 所以 $\mathbf{0}=\boldsymbol{\eta}_0=\Sigma b_i\boldsymbol{\beta}_i$ 和 $\mathbf{0}=\Sigma c_j\gamma_j$. 但是 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_k$ 线性无关, $\gamma_1,\cdots,\gamma_m$ 线性无关, 因此 $b_1=\cdots=b_k=0$, 和 $c_1=\cdots=c_m=0$. 于是关系式 (33) 只当所有系数为零时才成立, 因此 $\boldsymbol{\beta}_1\cdots,\boldsymbol{\beta}_k,\gamma_1,\cdots,\gamma_m$ 确实线性无关.

这个定理及其证明可以很容易地推广到有限多个子空间的直和的情形.

推论　*如果有限维向量空间 V 是它的子空间 S 和 T 的直和, 那么*

$$d[V]=d[S]+d[T]. \tag{34}$$

式中 $d[V]$ 表示空间 V 的维数, 等等.

证明　因为空间的维数等于 (任意) 基向量的个数, 所以上述证明表明, 如果 $d[S]=k, d[T]=m$, 则 $d[V]=k+m$.　证毕

当 V 是 S 和 T 的直和时, 我们称 S 和 T 是 V 的*补子空间*. 那么我们有

$$S+T=V,\quad S\cap T=\mathbf{0}. \tag{35}$$

事实上, (32) 式指出 V 是子空间 S 和 T 的线性和. 断言 $S\cap T=\mathbf{0}$ 证明如下, 如果 $\boldsymbol{\xi}_1$ 是 S 和 T 的任意公共向量, 那么 $\boldsymbol{\xi}$ 有形为 (32) 的两种表示 $\boldsymbol{\xi}_1=\boldsymbol{\xi}_1+\mathbf{0}$ 或 $\boldsymbol{\xi}_1=\mathbf{0}+\boldsymbol{\xi}_1$; 因为这两种表示一定是一样的, 所以 $\boldsymbol{\xi}_1=\mathbf{0}$, 因此交 $S\cap T$ 是零向量. 反过来我们可以证明, 如果条件 (35) 成立, 则 V 是 S 和 T 的直和. 于是, 在这种情况下, 关系式 (34) 可化为

$$d[V]=d[S+T]+d[S\cap T]=d[S]+d[T].$$

上面的后一个等式对任意两个子空间情形也成立。

定理 17 设 S 和 T 是向量空间 V 的任意两个有限维子空间, 那么

$$d[S]+d[T]=d[S\cap T]+d[S+T]. \tag{36}$$

证明 设 $\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_n$ 是 $S\cap T$ 的一组基, 根据定理 6, S 和 T 分别有基 $\xi_1,\cdots,\xi_n$, $\boldsymbol{\eta}_1,\cdots,\boldsymbol{\eta}_r$ 和 $\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_n$, $\boldsymbol{\zeta}_1,\cdots,\boldsymbol{\zeta}_s$. 显然 $\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_n$, $\boldsymbol{\eta}_1,\cdots,\boldsymbol{\eta}_r$, $\boldsymbol{\zeta}_1,\cdots,\boldsymbol{\zeta}_s$ 一起张成 $S+T$. 它们也是一组基, 这因为

$$a_1\boldsymbol{\xi}_1+\cdots+\alpha_n\boldsymbol{\xi}_n+b_1\boldsymbol{\eta}_1+\cdots+b_r\boldsymbol{\eta}_r+c_1\boldsymbol{\zeta}_1+\cdots+c_s\boldsymbol{\zeta}_s=\mathbf{0}$$

推出 $\Sigma b_j\boldsymbol{\eta}_j=-\Sigma a_i\boldsymbol{\xi}_i-\Sigma c_k\boldsymbol{\zeta}_k$ 在 T 中, 因此它在 $S\cap T$ 中, 所以 $\Sigma b_j\boldsymbol{\eta}_j=\Sigma d_i\boldsymbol{\xi}_i$, 其中 d_i 是某一组标量. 因为 $\boldsymbol{\xi}_i$ 和 $\boldsymbol{\eta}_j$ 线性无关, 因此每个 b_i 必为 0. 类似地, 每个 $c_k=0$, 代入原式得 $\Sigma a_i\boldsymbol{\xi}_i=\mathbf{0}$, 所以每个 $a_i=0$. 这就证明了 $\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_n$, $\boldsymbol{\eta}_1,\cdots,\boldsymbol{\eta}_r$, $\boldsymbol{\zeta}_1,\cdots,\boldsymbol{\zeta}_s$ 是 $S+T$ 的一组基.

证明了这个之后, 我们看到定理的结论归结为算术公式 $(n+r)+(n+s)=n+(n+r+s)$.

习 题

1. 在 7.6 节的习题 4 中, 指出哪些向量集合是包含它们的空间的基.
2. 在 $\mathbf{Q}^4$ 中求出单位向量 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4$ 关于基 $\boldsymbol{\alpha}_1=(1,1,0,0)$, $\boldsymbol{\alpha}_2=(0,0,1,1),\boldsymbol{\alpha}_3=(1,0,0,4),\boldsymbol{\alpha}_4=(0,0,0,2)$ 的坐标.
3. 求出向量 $(1,0,1)$ 关于 $\mathbf{C}^3$ 的基

   $$(2\,\mathrm{i},1,0),\quad (2,-\mathrm{i},1),\quad (0,1+\mathrm{i},1-\mathrm{i})$$

   的坐标.
4. 在 $\mathbf{Q}^4$ 中求出
   (a) 包含向量 (1, 2, 1, 1) 的基;
   (b) 包含向量 (1, 1, 0, 2) 和 $(1,-1,2,0)$ 的基;
   (c) 包含向量 (1, 1, 0, 0), (0, 0, 2, 2), (0, 2, 3, 0) 的基.
5. 证明：以有理数 $a,\cdots,e$ 为系数的数

   $$a+b\sqrt{2}+c\sqrt{3}+d\sqrt{6}+e\sqrt{12}$$

   的全体构成一个交换环, 这个环是有理数域 $\mathbf{Q}$ 上的向量空间. 求出这个空间的一组基.
6. 在 $\mathbf{Q}^4$ 中, 两个子空间 S 和 T 分别由下面向量张成：

   $$S:\ (1,-1,2,-3),(1,1,2,0),(3,-1,6,-6),$$
   $$T:\ (0,-2,0,-3),(1,0,1,0).$$

   求出 $S,T,S\cap T$ 和 $S+T$ 的维数.

*7. 对一般的域 $\mathbf{Z}_p^4$ 解习题 6.

8. 设 S 和 T 是 F^n 中具有固定维数分别为 s 和 t 的可变的子空间, 求 $S+T$ 的最大可能维数和 $S\cap T$ 的最小可能维数. 证明你的结论.

*9. 证明: 对于子空间, 由 $S\cap T=S\cap T'$, $S+T=S+T'$ 和 $T\subset T'$, 可推出 $T=T'$.

10. 设 S 是有限维向量空间 V 的子空间, 证明: 存在 V 的一个子空间 T , 使得 V 是 S 和 T 的直和.

*11. 设 $S_1,\cdots,S_p$ 是 V 的子空间, 如果 V 的每个向量 $\boldsymbol{\xi}$ 具有唯一的表示 $\boldsymbol{\xi}=\boldsymbol{\eta}_1+\cdots+\boldsymbol{\eta}_p$, 其中 $\boldsymbol{\eta}_i\in S_i$, 则称 V 是 $S_1,\cdots,S_p$ 的直和. 对这样的直和叙述并证明与定理 10 相类似的定理.

12. 证明: V 是 S 和 T 的直和当且仅当 (35) 成立.

*13. 对 p 个子空间的直和叙述并证明与习题 12 相类似的定理.

14. 向量空间 V 的自同构指的是 V 同它自身的同构.

    (a) 证明: 对应 $(x_1,x_2,x_3)\longmapsto(x_2,-x_1,x_3)$ 是 F^3 的自同构.

    (b) 证明: V 的所有自同构组成的集合是 V 上的变换群.

15. F^2 的一个自同构把 (1, 0) 映射到 (0, 1), 把 (0, 1) 映射到 (−1, −1). 它的阶是多少? 你的答案依赖于基域吗?

*16. 建立有限维向量空间的自同构和它的有序基之间的一一对应 (参看习题 14). $\mathbf{Z}_2$ 有多少自同构? $\mathbf{Z}_p^n$ 呢?

7.9　内　　积

普通空间是实数域上的三维向量空间, 它记作 $\mathbf{R}^3$. 在这个空间中可以用公式来定义向量的长度和向量之间的夹角(包括直角), 这些公式不仅顺利地推广到 $\mathbf{R}^n$ 空间, 而且还推广到*无穷维*实向量空间 (见 7.10 节的例 2). 这些推广将是 7.9 节至 7.11 节中的课题.

为了建立有关的公式, 我们需要另外一种运算. 为此目的, 最方便的是做内积的运算. 含有实分量的两个向量 $\boldsymbol{\xi}=(x_1,\cdots,x_n)$ 和 $\boldsymbol{\eta}=(y_1,\cdots,y_n)$ 的*内积*指的是数量

$$(\boldsymbol{\xi},\boldsymbol{\eta})=x_1y_1+x_2y_2+\cdots+x_ny_n. \tag{37}$$

(因为这是一个标量, 所以物理学家常常把我们上面的内积说成两个向量的"标量积".) 内积有四个重要性质, 这些性质都是定义 (37) 的直接结论:

$$(\boldsymbol{\xi}+\boldsymbol{\eta},\boldsymbol{\zeta})=(\boldsymbol{\xi},\boldsymbol{\zeta})+(\boldsymbol{\eta},\boldsymbol{\zeta}),\qquad(c\boldsymbol{\xi},\boldsymbol{\eta})=c(\boldsymbol{\xi},\boldsymbol{\eta}); \tag{38}$$

$$(\boldsymbol{\xi},\boldsymbol{\eta})=(\boldsymbol{\eta},\boldsymbol{\xi}),\qquad(\boldsymbol{\xi},\boldsymbol{\xi})>0\text{ 除非 }\boldsymbol{\xi}=\mathbf{0}. \tag{39}$$

前两个定律表明内积对于左边因子是*线性的*; 第三个定律是*对称律*, 因此同前两个公式一起得出, 内积对于左右因子都是线性的 (*双线性*); 第四个定律是*正性律*.

例如, 计算平面 $\mathbf{R}^2$ 中向量 $\boldsymbol{\xi}$ 的长度$|\boldsymbol{\xi}|$(也称为"绝对值"或"模")的笛卡儿公式给出这个长度是内积的平方根

$$|\boldsymbol{\xi}| = \sqrt{x_1^2 + x_2^2} = (\boldsymbol{\xi}, \boldsymbol{\xi})^{\frac{1}{2}}. \tag{40}$$

三维空间中的长度用一个类似的公式来计算. 再有, 如果 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 是任意两个向量, 那么对于以 $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma} = \boldsymbol{\beta} - \boldsymbol{\alpha}$ 为三边的三角形 (图 7-2), 由三角余弦定理得到

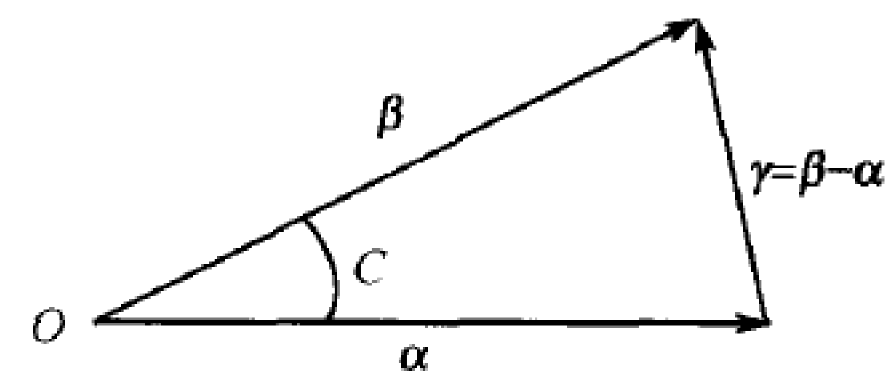


图　7-2

$$|\boldsymbol{\beta} - \boldsymbol{\alpha}|^2 = |\boldsymbol{\alpha}|^2 + |\boldsymbol{\beta}|^2 - 2|\boldsymbol{\alpha}| \cdot |\boldsymbol{\beta}| \cdot \cos C,$$

$(C = \angle(\boldsymbol{\alpha}, \boldsymbol{\beta}))$. 而根据 (38) 和 (40) 有

$$|\boldsymbol{\beta} - \boldsymbol{\alpha}|^2 = (\boldsymbol{\beta} - \boldsymbol{\alpha}, \boldsymbol{\beta} - \boldsymbol{\alpha}) = (\boldsymbol{\beta}, \boldsymbol{\beta}) - 2(\boldsymbol{\alpha}, \boldsymbol{\beta}) + (\boldsymbol{\alpha}, \boldsymbol{\alpha}).$$

与上式合并并消去一些项, 我们得到

$$\cos\angle(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{(\boldsymbol{\alpha}, \boldsymbol{\beta})}{|\boldsymbol{\alpha}||\boldsymbol{\beta}|}. \tag{41}$$

也就是说, 两个向量 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 之间的夹角 $\angle(\boldsymbol{\alpha}, \boldsymbol{\beta})$ 的余弦是这两个向量的内积与它们长度之积的比. 由这个公式可以得到, 几何上两个向量 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$*正交*(或垂直) 当且仅当内积 $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ 为零.

由于向量加法和数乘运算容易推广到任意域上的任意维空间中去, 自然希望把长度和角度的概念做类似的推广. 然而, 当我们这样推广时却发现, 虽然维数可以是任意的, 但是推广到很多数域时产生了麻烦. 即使内积可以由 (37) 定义, 但是长度

$$|\boldsymbol{\xi}| = (\boldsymbol{\xi}, \boldsymbol{\xi})^{\frac{1}{2}} = (x_1^2 + x_2^2 + \cdots + x_n^2)^{\frac{1}{2}} \tag{42}$$

是没有定义的, 除非每个 n 平方和有平方根. 对于距离也有同样问题, 而角度的推广引起更多的困难.

由于这些原因, 我们现在只限于讨论实数域上向量空间中的长度、角度和有关课题. 在 9.12 节中我们将讨论复数域上相应的概念.

习　题

1. 用解析几何方法证明: 在平面上, 向量 $\boldsymbol{\xi} = (x_1, x_2)$ 和 $\boldsymbol{\eta} = (y_1, y_2)$ 之间的距离平方等于 $|\boldsymbol{\xi}|^2 + |\boldsymbol{\eta}|^2 - 2(\boldsymbol{\xi}, \boldsymbol{\eta})$.

2. 利用三维空间中的方向余弦证明：两个向量 $\boldsymbol{\xi}$ 和 $\boldsymbol{\eta}$ 正交当且仅当 $(\boldsymbol{\xi},\boldsymbol{\eta})=0$.
3. 如果复分量向量 $\boldsymbol{\xi}$ 的长度是由公式 (42) 定义的, 证明：存在长度为零的非零向量.
4. 证明：在域 $\mathbf{Z}_3$ 和 $\mathbf{Q}$ 中存在一个二平方和, 它没有平方根.
5. 从定义 (37), 证明公式 (38) 和 (39).
6. 证明类似于 (38) 的公式, 这个公式断言：内积对于右边因子是线性的.
7. 证明：任意平行四边形对角线长度的平方和等于它四个边长的平方和.

*8. 在 $\mathbf{R}^3$ 中用

$$\boldsymbol{\xi}\times\boldsymbol{\eta}=(x_2y_3-x_3y_2,x_3y_1-x_1y_3,x_1y_2-x_2y_1)$$

定义两个向量 $\boldsymbol{\xi}$ 和 $\boldsymbol{\eta}$ 的外积.

(a) 证明：$(\boldsymbol{\xi}\times\boldsymbol{\eta},\boldsymbol{\zeta}\times\boldsymbol{\tau})=(\boldsymbol{\xi},\boldsymbol{\zeta})(\boldsymbol{\eta},\boldsymbol{\tau})-(\boldsymbol{\xi},\boldsymbol{\tau})(\boldsymbol{\eta},\boldsymbol{\zeta})$.

(b) 设 $\boldsymbol{\xi}=\boldsymbol{\zeta},\boldsymbol{\eta}=\boldsymbol{\tau}$, 推导 $\mathbf{R}^3$ 中的施瓦兹 (Schwarz) 不等式作为 (a) 的推论. (参看定理 18.)

(c) 证明 $\boldsymbol{\xi}\times(\boldsymbol{\eta}\times\boldsymbol{\zeta})=(\boldsymbol{\xi},\boldsymbol{\zeta})\boldsymbol{\eta}-(\boldsymbol{\xi},\boldsymbol{\eta})\boldsymbol{\zeta}$.

7.10 欧几里得向量空间

对维数不加限制的几何讨论是建立在下面定义的基础上, 这个定义是根据 7.9 节的考虑而提出来的.

定义 *一个具有实标量的向量空间 E, 如果 E 中任意两个向量 $\boldsymbol{\xi}$ 和 $\boldsymbol{\eta}$ 对应一个 (实的) 内积 $(\boldsymbol{\xi},\boldsymbol{\eta})$, 它在 (38) 和 (39) 意义下具有对称性、双线性和正性, 那么 E 称为欧几里得向量空间.*

例 1 任意 $\mathbf{R}^n$, 如果其中 $(\boldsymbol{\xi},\boldsymbol{\eta})$ 是由 (37) 式定义, 那么它是 n 维欧几里得向量空间.

例 2 定义在区间 $0\leqslant x\leqslant 1$ 上的全体连续实函数 $\phi(x)$, 如果我们定义内积 $(\phi,\psi)=\int_0^1\phi(x)\psi(x)\mathrm{d}x$, 那么它构成一个无穷维欧几里得向量空间.

欧几里得向量空间 E 的向量 $\boldsymbol{\xi}$ 的长度 $|\boldsymbol{\xi}|$ 可以定义为内积的平方根 $(\boldsymbol{\xi},\boldsymbol{\xi})^{\frac{1}{2}}$——(39) 的正性条件保证了平方根的存在.

定理 18 *在任意欧几里得向量空间中, 长度具有下列性质：*

(i) $|c\boldsymbol{\xi}|=|c|\cdot|\boldsymbol{\xi}|$;

(ii) $|\boldsymbol{\xi}|>0$, 除非 $\boldsymbol{\xi}=\mathbf{0}$;

(iii) $|(\boldsymbol{\xi},\boldsymbol{\eta})|\leqslant|\boldsymbol{\xi}|\cdot|\boldsymbol{\eta}|$ (施瓦兹不等式);

(iv) $|\boldsymbol{\xi}+\boldsymbol{\eta}|\leqslant|\boldsymbol{\xi}|+|\boldsymbol{\eta}|$ (三角形不等式).

证明 因为 $(c\boldsymbol{\xi},c\boldsymbol{\xi})=c^2(\boldsymbol{\xi},\boldsymbol{\xi})$, 所以我们有性质 (i). 性质 (ii) 是欧几里得向量空间的定义中所要求的正性条件的推论.

性质 (iii) 的证明并不直接. 如果 $\boldsymbol{\xi}=\mathbf{0}$ 或者 $\boldsymbol{\eta}=\mathbf{0}$, 那么 (iii) 归结为平凡的不等式 $0\leqslant 0$. 否则,

$$0\leqslant(a\boldsymbol{\xi}\pm b\boldsymbol{\eta},a\boldsymbol{\xi}\pm b\boldsymbol{\eta})=a^2(\boldsymbol{\xi},\boldsymbol{\xi})\pm 2ab(\boldsymbol{\xi},\boldsymbol{\eta})+b^2(\boldsymbol{\eta},\boldsymbol{\eta}).$$

令 $a=|\boldsymbol{\eta}|,b=|\boldsymbol{\xi}|$, 故 $a^2=(\boldsymbol{\eta},\boldsymbol{\eta}),b^2=(\boldsymbol{\xi},\boldsymbol{\xi})$. 代入上式则有

$$\mp 2|\boldsymbol{\xi}|\cdot|\boldsymbol{\eta}|\cdot(\boldsymbol{\xi},\boldsymbol{\eta})\leqslant 2(\boldsymbol{\xi},\boldsymbol{\xi})(\boldsymbol{\eta},\boldsymbol{\eta})=2|\boldsymbol{\xi}|^2\cdot|\boldsymbol{\eta}|^2. \tag{43}$$

两边除以 $2|\boldsymbol{\xi}|\cdot|\boldsymbol{\eta}|>0$, 我们就得性质 (iii).

由 (iii) 容易得到性质 (iv), 这因为

$$\begin{aligned}|\boldsymbol{\xi}+\boldsymbol{\eta}|^2&=(\boldsymbol{\xi}+\boldsymbol{\eta},\boldsymbol{\xi}+\boldsymbol{\eta})=(\boldsymbol{\xi},\boldsymbol{\xi})+2(\boldsymbol{\xi},\boldsymbol{\eta})+(\boldsymbol{\eta},\boldsymbol{\eta})\\&\leqslant|\boldsymbol{\xi}|^2+2|\boldsymbol{\xi}|\cdot|\boldsymbol{\eta}|+|\boldsymbol{\eta}|^2=(|\boldsymbol{\xi}|+|\boldsymbol{\eta}|)^2.\end{aligned}$$

现在, 如果我们定义 E 中任意两个向量 $\boldsymbol{\xi}$ 和 $\boldsymbol{\eta}$ 之间的*距离*为 $|\boldsymbol{\xi}-\boldsymbol{\eta}|$, 则我们可以证明它具有普通距离的所谓"度量"性质, 首先是由弗雷谢 (Fréchet, 1906) 做了抽象的考虑.

定理 19 *距离具有性质:*

$(M1)$ $|\boldsymbol{\xi}-\boldsymbol{\xi}|=0$, 而当 $\boldsymbol{\xi}\neq\boldsymbol{\eta}$, $|\boldsymbol{\xi}-\boldsymbol{\eta}|>0$;

$(M2)$ $|\boldsymbol{\xi}-\boldsymbol{\eta}|=|\boldsymbol{\eta}-\boldsymbol{\xi}|$ (对称性);

$(M3)$ $|\boldsymbol{\xi}-\boldsymbol{\eta}|+|\boldsymbol{\eta}-\boldsymbol{\zeta}|\geqslant|\boldsymbol{\xi}-\boldsymbol{\zeta}|$.

证明 首先, 根据性质 (i) 有 $|\boldsymbol{\xi}-\boldsymbol{\xi}|=|\mathbf{0}|=|0\cdot\boldsymbol{\xi}|=0\cdot|\boldsymbol{\xi}|=0$, 而根据 (ii), 当 $\boldsymbol{\xi}-\boldsymbol{\eta}\neq\mathbf{0}$(或 $\boldsymbol{\xi}\neq\boldsymbol{\eta}$), 有 $|\boldsymbol{\xi}-\boldsymbol{\eta}|>0$, 这就证明了 $(M1)$. 其次, 根据性质 (i) 有 $|\boldsymbol{\xi}-\boldsymbol{\eta}|=|(-1)(\boldsymbol{\eta}-\boldsymbol{\xi})|=|-1|\cdot|\boldsymbol{\eta}-\boldsymbol{\xi}|=|\boldsymbol{\eta}-\boldsymbol{\xi}|$, 这就证明了 $(M2)$. 最后, $(M3)$ 由 (iv) 推出, 这因为

$$|\boldsymbol{\xi}-\boldsymbol{\eta}|+|\boldsymbol{\eta}-\boldsymbol{\zeta}|\geqslant|(\boldsymbol{\xi}-\boldsymbol{\eta})+(\boldsymbol{\eta}-\boldsymbol{\zeta})|=|\boldsymbol{\xi}-\boldsymbol{\zeta}|.$$

从施瓦兹不等式我们特别推出, 对任意非零向量 $\boldsymbol{\xi},\boldsymbol{\eta}$, 有 $-1\leqslant\dfrac{(\boldsymbol{\xi},\boldsymbol{\eta})}{|\boldsymbol{\xi}|\cdot|\boldsymbol{\eta}|}\leqslant 1$. 因此在 $0°$和$180°$ 之间有一个且仅有一个角, 它的余弦是 $\dfrac{(\boldsymbol{\xi},\boldsymbol{\eta})}{|\boldsymbol{\xi}|\cdot|\boldsymbol{\eta}|}$, 我们就可以把这个角定义为向量 $\boldsymbol{\xi}$ 和 $\boldsymbol{\eta}$ 之间的*夹角* (同 (41) 的特殊情况相比较). 除了直角的情形外, 我们不去证明如此定义的角有什么性质 (你能证明 $\angle(\boldsymbol{\xi},\boldsymbol{\eta})+\angle(\boldsymbol{\eta},\boldsymbol{\zeta})\geqslant\angle(\boldsymbol{\xi},\boldsymbol{\zeta})$ 吗?)

两个向量 ξ 和 η, 如果满足 $(\boldsymbol{\xi},\boldsymbol{\eta})=0$, 则称它们是*正交的* (记作 $\boldsymbol{\xi}\perp\boldsymbol{\eta}$). 把这个定义用到上面的例 2 中, 就得到分析上一个重要概念, 即正交函数的概念. 容易证明, 如果 $\boldsymbol{\xi}\perp\boldsymbol{\eta}$, 则 $\boldsymbol{\eta}\perp\boldsymbol{\xi}$(正交关系是对称的), 并且对一切标量 c 和 c' 有 $c\boldsymbol{\xi}\perp c'\boldsymbol{\eta}$. 还有, $\mathbf{0}$ 是唯一的与自身正交的向量. 此外, 如果 $(\boldsymbol{\eta},\boldsymbol{\xi}_1)=\cdots=(\boldsymbol{\eta},\boldsymbol{\xi}_m)=0$, 则对

任意标量 $c_1, \cdots, c_m$, 有

$$(\boldsymbol{\eta}, c_1\boldsymbol{\xi}_1 + \cdots + c_m\boldsymbol{\xi}_m) = c_1(\boldsymbol{\eta}, \boldsymbol{\xi}_1) + \cdots + c_m(\boldsymbol{\eta}, \boldsymbol{\xi}_m)$$
$$= c_1 \cdot 0 + \cdots + c_m \cdot 0 = 0,$$

所以 $\boldsymbol{\eta}$ 与 $\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_m$ 的每个线性组合也是正交的. 这就证明了

定理 20　如果一个向量与 $\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_m$ 正交, 那么它与由 $\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_m$ 张成的空间中每个向量正交.

习　题

1. 设 $\boldsymbol{\xi} = (1, 2, 3, 4), \boldsymbol{\eta} = (0, 3, -2, 1)$, 计算 $(\boldsymbol{\xi}, \boldsymbol{\eta}), |\boldsymbol{\xi}|, |\boldsymbol{\eta}|, \angle(\boldsymbol{\xi}, \boldsymbol{\eta})$.
2. 设 $\boldsymbol{\xi}$ 和 $\boldsymbol{\eta}$ 如习题 1 所述, 求出形为 $(1, 1, 0, 0) + c_1\boldsymbol{\xi} + c_2\boldsymbol{\eta}$ 并与 $\boldsymbol{\xi}$ 和 $\boldsymbol{\eta}$ 两个向量正交的向量.
3. (a) 在正文的例 2 中, $\sin 2\pi x$ 与 $\cos 2\pi x$ 正交吗?

   (b) $\sin 2m\pi x$ 与 $\sin 2n\pi x$ 正交吗?

   (c) 求出与 1 和 x 正交的二次多项式.
4. 证明: $|\boldsymbol{\xi} - \boldsymbol{\eta}|^2 + |\boldsymbol{\xi} + \boldsymbol{\eta}|^2 = 2(|\boldsymbol{\xi}|^2 + |\boldsymbol{\eta}|^2)$.
5. 证明: 在 $\mathbf{R}^3$ 中, 恰好存在两个长度为 1 的向量同两个已知线性无关向量垂直.
6. 证明: 在 $\mathbf{R}^3$ 中, 存在一个含有有理坐标的向量与任意两个给定的含有有理坐标的向量垂直.
7. 设 $\boldsymbol{\alpha}, \boldsymbol{\beta} \neq \mathbf{0}$ 是欧几里得向量空间的两个固定向量, 求出形为 $\boldsymbol{\gamma} = \boldsymbol{\alpha} + t\boldsymbol{\beta}$ 的最短向量. 这个向量与 $\boldsymbol{\beta}$ 正交吗? 画出图来.

*8. 证明: 如果向量 $\boldsymbol{\alpha}$ 到 $\boldsymbol{\beta}$ 的距离同到 $\boldsymbol{\gamma}$ 的距离相等, 那么线段 $\overline{\boldsymbol{\beta}\boldsymbol{\gamma}}$ 的中点是从 $\boldsymbol{\alpha}$ 到 $\overline{\boldsymbol{\beta}\boldsymbol{\gamma}}$ 的垂线的垂足.

9. 证明: 在欧几里得向量空间中, 如果 $|\boldsymbol{\xi}| = |\boldsymbol{\alpha}|$, 那么 $\boldsymbol{\xi} - \boldsymbol{\alpha} \perp \boldsymbol{\xi} + \boldsymbol{\alpha}$. 从几何上解释这个结论.

*10. (a) 证明: 二次方程

$$(\boldsymbol{\xi}, \boldsymbol{\xi})t^2 + 2(\boldsymbol{\xi}, \boldsymbol{\eta})t + (\boldsymbol{\eta}, \boldsymbol{\eta}) = |t\boldsymbol{\xi} + \boldsymbol{\eta}|^2 = 0$$

的判别式 $B^2 - 4AC$ 是 $4[(\boldsymbol{\xi}, \boldsymbol{\eta})^2 - (\boldsymbol{\xi}, \boldsymbol{\xi})(\boldsymbol{\eta}, \boldsymbol{\eta})]$.

(b) 利用上述事实证明施瓦兹不等式.

(提示: $|t\boldsymbol{\xi} + \boldsymbol{\eta}| = 0$ 不可能有两个不同的实解 t, 除非 $\boldsymbol{\xi} = \mathbf{0}$.)

11. 证明: 在任意欧几里得向量空间中, 有 $||\boldsymbol{\xi}| - |\boldsymbol{\eta}|| \leqslant |\boldsymbol{\xi} - \boldsymbol{\eta}|$.
12. 证明: 如果 $\mathbf{R}^3$ 的内积是由

$$(\boldsymbol{\xi}, \boldsymbol{\eta}) = (x_1 + x_2)(y_1 + y_2) + x_2y_2 + (x_2 + 2x_3)(y_2 + 2y_3)$$

来定义的, 那么 $\mathbf{R}^3$ 就成为欧几里得向量空间.

7.11 标准正交基

在 7.10 节的例 1 中, 单位向量 $\varepsilon_1=(1,0,\cdots,0),\cdots,\varepsilon_n=(0,0,\cdots,1)$ 具有单位长度并且互相正交. 这是"标准正交基"的一个例子.

定义 一组向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 满足下列条件时称为标准正交的:

(i) 对一切 i, 有 $|\boldsymbol{\alpha}_i|=1$; (ii) 当 $i\neq j$, 有 $\boldsymbol{\alpha}_i\perp\boldsymbol{\alpha}_j$.

引理 1 欧几里得向量空间 E 的一组非零正交向量 $\alpha_1,\cdots,\alpha_n$ 线性无关.

证明 如果 $x_1\boldsymbol{\alpha}_1+\cdots+x_m\boldsymbol{\alpha}_m=\mathbf{0}$, 则对 $k=1,\cdots,m$ 有

$$0=(\mathbf{0},\boldsymbol{\alpha}_k)=x_1(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_k)+\cdots+x_m(\boldsymbol{\alpha}_m,\boldsymbol{\alpha}_k)=x_k(\boldsymbol{\alpha}_k,\boldsymbol{\alpha}_k),$$

这里最后一个等式是从正交性的假设得来的. 但是根据假设 $\boldsymbol{\alpha}_k\neq\mathbf{0}$, 因此 $(\boldsymbol{\alpha}_k,\boldsymbol{\alpha}_k)>0$, 所以 $x_k=0$ 证毕

推论 张成空间 E 的标准正交向量组是 E 的一组基 (即所谓"标准正交基").

我们现在将指出, 欧几里得向量空间的任意一组基, 如何只通过有理运算把它正交化. 这称为格拉姆–施密特 (Gram-Schmidt)正交化方法.

引理 2 由有限维欧几里得向量空间 E 的任意一组 (有限个) 线性无关向量 $\gamma_1,\cdots,\gamma_m$, 可以构造一组非零正交向量

$$\boldsymbol{\alpha}_i=\boldsymbol{\gamma}_i-\sum_{k<i}d_{ik}\boldsymbol{\gamma}_k\qquad(i=1,\cdots,m),\tag{44}$$

它们同 $\gamma_1,\cdots,\gamma_m$ 张成 E 的相同的子空间.

证明 对 m 用归纳法, 我们可以假定非零正交向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{m-1}$ 已经构成, 它们同 $\gamma_1,\cdots,\gamma_{m-1}$ 张成相同的子空间 S. 我们现在把 γ_m 分成两部分:"平行"于 S 的部分 $\boldsymbol{\beta}_m$ 和垂直于 S 的部分 $\boldsymbol{\alpha}_m$. 为做到这一点, 令

$$\boldsymbol{\alpha}_m=\boldsymbol{\gamma}_m-\sum_{k<m}c_{mk}\boldsymbol{\alpha}_k,\quad \text{其中}\quad c_{mk}=\frac{(\boldsymbol{\gamma}_m,\boldsymbol{\alpha}_k)}{(\boldsymbol{\alpha}_k,\boldsymbol{\alpha}_k)},\tag{44'}$$

那么对 $j=1,\cdots,m-1$, 我们有

$$(\boldsymbol{\alpha}_m,\boldsymbol{\alpha}_j)=(\boldsymbol{\gamma}_m,\boldsymbol{\alpha}_j)-\sum_{k=1}^{m-1}c_{mk}(\boldsymbol{\alpha}_k,\boldsymbol{\alpha}_j)=0,$$

这是因为由正交性, 当 $k\neq j$ 时 $(\boldsymbol{\alpha}_k,\boldsymbol{\alpha}_j)=0$, 而由 (44′) 有 $c_{mj}(\boldsymbol{\alpha}_j,\boldsymbol{\alpha}_j)=(\boldsymbol{\gamma}_m,\boldsymbol{\alpha}_j)$. 根据归纳法假定, (44) 式中 $\boldsymbol{\alpha}_i(i=1,\cdots,m-1)$ 的表达式代入 (44′) 式中, 有

$$\boldsymbol{\alpha}_m=\boldsymbol{\gamma}_m-\sum_{k<m}c_{mk}\boldsymbol{\alpha}_k=\boldsymbol{\gamma}_m-\sum_{k<m}c_{mk}\boldsymbol{\gamma}_k+\sum_{j<k<m}c_{mk}d_{kj}\boldsymbol{\gamma}_j.$$

这就证明了 (44) 式对于 m 也成立, 其中

$$d_{mk} = c_{mk} - \sum_{k<j<m} c_{mj} d_{jk}.$$

因为 γ_m 与 $\gamma_1, \cdots, \gamma_{m-1}$ 线性无关, 所以它不可能在 S 中, 因此 $\boldsymbol{\alpha}_m \neq \mathbf{0}$. 最后, $\gamma_1, \cdots, \gamma_m$ 和 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 两组向量都张成由 S 和 γ_m 张成的子空间. 这就完成了引理 2 的证明.

定理 21　*有限维欧几里得向量空间 E 的每组标准正交向量 $\gamma_1, \cdots, \gamma_m$ 是 E 的标准正交基的一部分.*

证明　根据定理 6, $\gamma_1, \cdots, \gamma_m$ 是 E 的基 $\gamma_1, \cdots, \gamma_n$ 的一部分. 这组基可由引理 2 正交化, 然后再设 $\boldsymbol{\beta}_i = \dfrac{\boldsymbol{\alpha}_i}{|\boldsymbol{\alpha}_i|}$, 使它标准化; 而这个过程对原来的正交向量 $\gamma_1, \cdots, \gamma_m$ 没有任何变化.

推论　*任意有限维欧几里得向量空间 E 有标准正交基.*

格拉姆-施密特正交化方法还有其他涵义, 例如, 设 S 是欧几里得向量空间 E 的任意 m 维子空间, 如上所述, S 具有标准正交基 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$. 如果 γ 是不在 S 中的任意向量, 那么用上述正交化过程可以把 γ 表示成两个向量的和 $\gamma = \boldsymbol{\alpha} + \boldsymbol{\beta}$, 其中分量 $\boldsymbol{\beta}$ 在 S 中, 分量 $\boldsymbol{\alpha}$ 与 S 的每个向量垂直. 向量 $\boldsymbol{\beta}$ 称为 γ 在 S 上的*正 (交) 投影*.

本节最后, 我们来确定已知 (实) 有限维向量空间 V 上的全部内积. 显然, 如果 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 是 V 的任意一组基, 那么对任意向量 $\boldsymbol{\xi} = x_1\boldsymbol{\alpha}_1 + \cdots + x_n\boldsymbol{\alpha}_n$ 和 $\boldsymbol{\eta} = y_1\boldsymbol{\alpha}_1 + \cdots + y_n\boldsymbol{\alpha}_n$, 根据双线性我们有

$$(\boldsymbol{\xi}, \boldsymbol{\eta}) = \left(\sum x_i\boldsymbol{\alpha}_i, \sum y_k\boldsymbol{\alpha}_k\right) = \sum_{i,k} x_i y_k (\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_k). \tag{45}$$

于是, 任意两个向量的内积通过 n^2 个实常数 $(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_k) = a_{ik}$ 被确定为坐标 x_i 和 y_k 的某一个双线性型 $\sum\limits_{i,k} a_{ik}x_iy_k$. 因为 $(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_k) = (\boldsymbol{\alpha}_k, \boldsymbol{\alpha}_i)$, 所以这个形式是对称的.

反过来, F^n 中任意对称双线性型 $\sum\limits_{i,k} a_{ik}x_iy_k (a_{ik} = a_{ki})$ 满足 (38) 和 (39) 式的前三个条件. 第四个条件表明二次型 $\Sigma a_{ik}x_ix_k$ 是"正定的", 也就是说, $\Sigma a_{ik}x_ix_k > 0$, 除非每个 $x_i = 0$. 一个方阵是否正定的判别方法将在 9.9 节中推导.

对于标准正交基, 我们有 $(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_k) = 0$, 当 $i \neq k$; $(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_i) = 1$, 因此 (45) 化为

$$(\boldsymbol{\xi}, \boldsymbol{\eta}) = \sum_{i=1}^{n} x_i y_j = x_1 y_1 + \cdots + x_n y_n. \tag{46}$$

由这个公式我们可得出结论

定理 22 对于标准正交基, "抽象的" 内积表现为 "具体的" 形式 (46).

这样每个有限维欧几里得向量空间就同构于某个 $\mathbf{R}^n$.

习 题

1. 对下列各组向量张成的四维欧几里得向量空间的子空间, 求出标准正交基:
   (a) (1, 1, 0, 0,), (0, 1, 2, 0) 和 (0, 0, 3, 4);
   (b) (2, 0, 0, 0), (1, 3, 3, 0) 和 (0, 4, 6, 1).
   (提示: 先找出正交基, 然后再标准化.)
2. 画图说明向量在一维子空间上的正投影.
3. 求向量 $\boldsymbol{\beta} = (2,1,3)$ 在由 $\boldsymbol{\alpha} = (1,0,1)$ 张成的子空间上的正投影.
4. 求 $\boldsymbol{\beta} = (0,0,0,3)$ 在习题 1 中所述的每个子空间上的正投影.
5. 设 S 是欧几里得向量空间 E 的任意子空间, 证明: 与 S 中每个向量 $\boldsymbol{\xi}$ 正交的全体向量的集合 $S^{\perp}$ 是满足下面条件的子空间:

$$S \cap S^{\perp} = \mathbf{0},\ S + S^{\perp} = E,\ 并且 d[S] + d[S^{\perp}] = d[E]$$

   (子空间 $S^{\perp}$ 称为 S 的正交补空间.)
6. 在三维欧几里得向量空间中, 求出由 $(2,-1,-2)$ 张成的子空间的正交补空间的基.
7. 求出习题 1 中所述的每个子空间的正交补空间的基.

*8. (a) 列出 $\mathbf{Q}^3$ 中的非平凡子空间, 它不包含任何长度为 1 的向量.
   (b) 对标量属于任意有序域的向量空间, 叙述并证明类似于引理 2 的命题.

7.12 商 空 间

我们现在将要指出, 6.13 节中的商群的构造方法容易推广到向量空间中去. 设 V 是域 F 上任意向量空间, 并设 S 是 V 的任意子空间. 在加法运算之下, V 是一个交换群, 而且 S 是 V 的一个 (正规) 子群. 因此我们可以构造加法商群 V/S.

例如, 在欧几里得空间 $\mathbf{R}^3$ 中, 设 S 是由单位向量 $(0,1,0)$ 的全体倍数 $(0,y,0)$ 组成. 则对任意向量 $\boldsymbol{\alpha} = (a,b,c)$, 陪集是由全体向量 $(a,b+y,c)$ 组成, 其中每个向量与 $\boldsymbol{\alpha}$ 有相同的 x 坐标 a 和相同的 z 坐标 c; 它们是向量 $(a,\cdot,c)$, 这里的圆点位置是一个任意元素. 在商群 $\mathbf{R}^3/S$ 中, 两个这样的向量的和 $(a,\cdot,c)+(a',\cdot,c')$ 显然是 $(a+a',\cdot,c+c')$.

在这个例子中, 我们也可以用任意标量 $t \in \mathbf{R}$ 去乘每个向量 $(a,\cdot,c)$ 而得到新的陪集 $(ta,\cdot,tc)$. 显然, 在这些运算之下商群 $\mathbf{R}^3/S$ 是一个 (实)向量空间. 我们现在指出, 类似的构造能够推广到一般情形.

已知域 F 上向量空间 V, 我们可以把 6.13 节的讨论移植到 V 上得到商空间$V/S = X$. 回忆一下, 对任意群 G 和 (正规) 子群 N , 商群 G/N 的元素只不过

是 N 在 G 中的陪集 xN. 因此, 已知向量空间 V 的一个子空间 S, 每个向量 $\boldsymbol{\alpha} \in V$ 确定 S 的一个陪集, 这个陪集定义为由所有和 $\boldsymbol{\alpha} + \boldsymbol{\sigma}$(对一切 $\boldsymbol{\sigma} \in S$) 组成的集合 $\boldsymbol{\alpha} + S$. 例如, $\boldsymbol{\alpha} = \boldsymbol{\alpha} + \mathbf{0}$ 是这个陪集的一个向量, 称它是这个陪集的"代表元素". 两个陪集 $\boldsymbol{\alpha} + S$ 和 $\boldsymbol{\beta} + S$ 相等 (作为集合) 当且仅当 $\boldsymbol{\alpha} - \boldsymbol{\beta} \in S$; 当这个结论成立时, $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 代表同一个陪集 (是陪集的元素). 几何上, 子空间 S 的不同陪集恰恰是 S 在平移之下的"平行子空间".

我们定义两个陪集的和是一个陪集:

$$(\boldsymbol{\alpha} + S) + (\boldsymbol{\beta} + S) = (\boldsymbol{\alpha} + \boldsymbol{\beta}) + S,$$

同 6.13 节的引理 2 中所说的一样, 这个和不依赖于代表元素 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 的选择. 下面定义陪集 $\boldsymbol{\alpha} + S$ 用标量 c 乘而得到的积是陪集

$$c(\boldsymbol{\alpha} + S) = c\boldsymbol{\alpha} + S.$$

因为 $\boldsymbol{\alpha} - \boldsymbol{\beta} \in S$ 可推出 $c\boldsymbol{\alpha} - c\boldsymbol{\beta} \in S$, 所以这个积也不依赖于已知陪集的代表元素的选择. 不难验证, 这两个定义使得 S 在 V 中的所有陪集的集合 V/S 成为一个向量空间, 它称为 V 对于 S 的商空间. 此外, 如果函数 P 由 $\boldsymbol{\alpha}P = \boldsymbol{\alpha} + S$ 定义, 那么 P 是向量空间的一个满同态, 其同态核恰好是 S, 值域是整个 V/S. 这个函数 P 称为 V 到它的商空间上的标准投影; 于是我们证明了:

定理 23　已知向量空间 V 的任意子空间 S, 则存在一个商空间 $X = V/S$ 和一个满同态 $P : V \to X$, 同态核是 S, 它的值域是 X.

习　题

1. 设 S 是空间 $\mathbf{R}^3$ 中的一维子空间, 证明: S 的全体陪集是所有平行于 S 的直线.
2. 设 $V = F^3$, F 是任意域, S 是由 (1, 1, 0) 和 (1, 1, 1) 张成的子空间.
   (a) 证明: 两个向量 (x, y, z) 和 (x', y', z') 在 S 的同一个陪集里当且仅当 $x + y' = x' + y$.
   (b) 当 $F = \mathbf{R}$, 描述 S 和它的陪集的几何意义.
3. 证明: 如果 S 是 $V = F^n$ 中同构于 F^m 的一个子空间, 那么 V/S 与 F^{n-m} 同构.
4. 详细证明: 在正文所述运算之下, 向量空间 V 的任意子空间 S 的全体陪集构成一个向量空间.
5. 设 $V = \mathbf{R}[x]$ 是所有实多项式 $f(x)$ 构成的空间, 并设

$$\phi : f(x) \longmapsto \frac{1}{2}[f(x) + f(-x)].$$

   (a) 证明: ϕ 是向量空间的同态.
   (b) 描述它的核 S 和商空间 V/S.

7.13 线性函数与对偶空间

在初等代数中, 有限维向量空间 $V = F^n$ 的变向量 $\boldsymbol{\xi} = (x_1, \cdots, x_n)$ 的坐标 $x_1, \cdots, x_n$ 的 (齐次) "线性函数" 是特殊形式的多项式函数

$$f(\boldsymbol{\xi}) = \boldsymbol{\xi} f = c_1x_1 + \cdots + c_nx_n = x_1c_1 + \cdots + x_nc_n, \tag{47}$$

这里 $c_1, \cdots, c_n$ 是域 F 中的任意常数. 容易验证, 任意这样的函数 f 满足恒等式

$$(\boldsymbol{\xi} + \boldsymbol{\eta})f = \boldsymbol{\xi} f + \boldsymbol{\eta} f, \qquad (a\boldsymbol{\xi})f = a(\boldsymbol{\xi} f), \tag{48}$$

其中 $\boldsymbol{\xi}$, $\boldsymbol{\eta}$ 为 V 中任意向量; a 为 F 中任意标量.

恒等式 (48) 与公式 (47) 的定义相比有两个优点: (i) (48) 是函数内在的性质 (即它们不依赖于 V 的基的选择); (ii) (48) 可用于无穷维向量空间 (例如用于函数空间). 因此, 我们把任意域 F 上任意向量空间 V 上的*线性函数* f 定义为满足两个恒等式 (48) 的从 V 到 F 的函数.

在第一个恒等式中取 $\boldsymbol{\eta} = \mathbf{0}$, 立即看出, $\mathbf{0}f = 0$. 这两个恒等式可推出组合恒等式

$$(a\boldsymbol{\xi} + b\boldsymbol{\eta})f = a(\boldsymbol{\xi} f) + b(\boldsymbol{\eta} f), \quad \boldsymbol{\xi}, \boldsymbol{\eta} \in V, \quad a, b \in F. \tag{49}$$

反过来, 当取 $a = b = 1$ 时这个恒等式便给出 (48) 的第一个恒等式, 因此 $\mathbf{0}f = 0$, 并且当取 $b = 0$ 时, 得出 (48) 的第二个恒等式. 简单地说, 线性函数 f 是保持线性组合性质的函数.

刚才定义的 "线性函数" 概念实质上等价于 7.8 节中引进的 "坐标" 概念, 即定理 14 中的每个 x_i, 当 $\boldsymbol{\xi}$ 在 V 上变化时, 它是 $\boldsymbol{\xi}$ 的线性函数. 下面结果是定理 14 的对偶定理, 在某种意义上, 定理 14 更简短明确.

定理 24 *如果 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_n$ 是 F 上向量空间 V 的一组基, $c_1, \cdots, c_n$ 是 F 中的 n 个常数, 那么在 V 中有一个且仅有一个线性函数 f, 使得 $\boldsymbol{\beta}_i f = c_i$, $i = 1, \cdots, n$. 这个函数由公式*

$$(x_1\boldsymbol{\beta}_1 + \cdots + x_n\boldsymbol{\beta}_n)f = x_1c_1 + \cdots + x_nc_n \tag{50}$$

给出.

证明 对 n 用归纳法, 对任意适合 $\boldsymbol{\beta}_i f = c_i (i = 1, \cdots, n)$ 的线性函数 f, 从 (49) 可直接推出方程 (50). 反过来, 对 V 的任意一组基 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_n$, 根据定理 14, 每个 $\boldsymbol{\xi}$ 有唯一的表示 $\boldsymbol{\xi} = x_1\boldsymbol{\beta}_1 + \cdots + x_n\boldsymbol{\beta}_n$, 因此对 F 中任意常数 $c_1, \cdots, c_n$, 方程 (50) 定义一个单值函数. 这个函数是线性的, 这因为对任意 $\boldsymbol{\xi}$ 和 $\boldsymbol{\eta} = y_1\boldsymbol{\beta}_1 + \cdots + y_n\boldsymbol{\beta}_n$, 有

$$\begin{aligned}(a\boldsymbol{\xi} + b\boldsymbol{\eta})f &= \left[\sum (ax_i + by_i)\boldsymbol{\beta}_i\right] f = \sum (ax_i + by_i)c_i \\ &= a\sum x_ic_i + b\sum y_ic_i = a(\boldsymbol{\xi} f) + b(\boldsymbol{\eta} f),\end{aligned}$$

因此条件 (49) 满足.

推论 F^n 上的线性函数是由线性表达式 (47) 给出的函数.

事实上, (47) 式给出函数 f, 它在 F^n 中单位向量 ε_i 上取值 c_i. 于是每个线性函数由 (47) 中的 n 个系数 $(c_1, \cdots, c_n)$ 所确定; 这就表示全体线性函数本身构成一个向量空间.

对任意向量空间 V , 把两个线性函数 f 和 g 的和 $f+g$ 定义为由方程

$$\boldsymbol{\xi}(f+g) = \boldsymbol{\xi}f + \boldsymbol{\xi}g, \quad 对一切 \quad \boldsymbol{\xi} \in V \tag{51}$$

给出的函数, 把线性函数 f 和标量 c 的乘积 cf 定义为由方程

$$\boldsymbol{\xi}(fc) = (\boldsymbol{\xi}f)c, \quad 对一切 \quad \boldsymbol{\xi} \in V, c \in F \tag{52}$$

给出的函数. 我们容易验证 $f+g$ 与 fc 仍是 V 上的线性函数.

定理 25 设 V 是 F 上的向量空间, V^* 是 V 上所有线性函数的集合, 那么 V^* 在由 (51) 和 (52) 定义的 $f+g$ 和 fc 两个运算之下也是 F 上的向量空间.

这个 V 上线性函数向量空间 V^* 称为 V 的对偶空间或 V 的共轭空间. 在现代数学中这是一个基本概念.

为证明定理, 我们只须验证, 对于运算 $f+g$ 与 fc, 向量空间的那些公理都成立. 例如, 为了证明分配律 $(f+g)c = fc + gc$, 我们注意, 对任意 $\boldsymbol{\xi} \in V$, 根据定义 (51) 和 (52) 以及 V 中的分配律, 有

$$\begin{aligned} \boldsymbol{\xi}[(f+g)c] &= [\boldsymbol{\xi}(f+g)]c = [\boldsymbol{\xi}f + \boldsymbol{\xi}g]c \\ &= (\boldsymbol{\xi}f)c + (\boldsymbol{\xi}g)c = \boldsymbol{\xi}(fc) + \boldsymbol{\xi}(gc) = \boldsymbol{\xi}(fc+gc). \end{aligned} \tag{53}$$

这个方程表明, 函数 $(f+g)c$ 和 $fc+gc$ 对任意自变量 $\boldsymbol{\xi}$ 都有相同的值, 因此它们一定相等. 其他公理的证明类似.

推论 1 如果向量空间 V 有一组有限基 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_n$, 那么它的对偶空间 V^* 有一组基, 这组基是由 $(x_1\boldsymbol{\beta}_1 + \cdots + x_n\boldsymbol{\beta}_n)f_i = x_i (i = 1, \cdots, n)$ 定义的 n 个线性函数 $f_1, \cdots, f_n$ 组成. 这 n 个线性函数由公式

$$\boldsymbol{\beta}_i f_j = \begin{cases} 0, & 当 i \neq j, \\ 1, & 当 i = j, \end{cases} \quad i, j = 1, \cdots, n \tag{54}$$

唯一确定.

证明 对于 n 个已知标量 $c_1, \cdots, c_n$, 线性组合 $f_1c_1 + \cdots + f_nc_n$ 是一个线性函数; 根据 (54), 它在任意基向量 $\boldsymbol{\beta}_i$ 上的值是

$$\boldsymbol{\beta}_i \left(\sum_j f_j c_j \right) = \sum_j \boldsymbol{\beta}_i f_j c_j = c_i.$$

我们可以推出函数 $f_1,\cdots,f_n$ 在 V^* 中线性无关, 这因为如果 $f=f_1c_1+\cdots+f_nc_n=\mathbf{0}$, 那么对每个 $i,\boldsymbol{\beta}_if=0$, 因此 $c_1=c_2=\cdots=c_n=0$. 还可以推出 n 个线性函数 $f_1,\cdots,f_n$ 张成空间 V^*: 根据定理 24, 任意线性函数 f 是由它的值 $\boldsymbol{\beta}_if=c_i$ 所确定的, 因此 f 等于以这些 c_i 值为系数而构成的线性组合 $\sum\limits_j f_jc_j$.

这组基 $f_1,\cdots,f_n$ 称为已知基 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n$ 的对偶基.

推论 2 *n 维向量空间 V 的对偶空间 V^* 的维数同 V 一样, 也是 n.*

把 V 的每个向量 $\Sigma x_i\boldsymbol{\beta}_i$ 映到 V^* 的函数 Σf_ic_i 的变换 $T:V\to V^*$ 是 V 到 V^* 上的同构; 然而, 这个同构依赖于 V 的基的选择.

设 $\boldsymbol{\xi}$ 是 V 中向量, f 是对偶空间 V^* 的向量, 我们也可以把 f 在自变量 $\boldsymbol{\xi}$ 上的值写成对称的"内积"记号 $\boldsymbol{\xi}f=(\boldsymbol{\xi},f)$. 那么方程 (49) 变成

$$(a\boldsymbol{\xi}+b\boldsymbol{\eta},f)=a(\boldsymbol{\xi},f)+b(\boldsymbol{\eta},f),\tag{55}$$

而加法和数乘的定义 (51) 和 (52) 变成

$$(\boldsymbol{\xi},fc+gd)=(\boldsymbol{\xi},f)c+(\boldsymbol{\xi},g)d.\tag{56}$$

这两个方程的类似暗示了另外一种解释. 在 $(\boldsymbol{\xi},f)$ 中, 保持 $\boldsymbol{\xi}$ 固定, 而让 f 变化. 那么, 由 (56), $\boldsymbol{\xi}$ 确定 f 的一个线性函数, 并且由 (55), 这些函数上的向量运算恰恰对应于原来向量 $\boldsymbol{\xi}$ 上的向量运算.

正式地, V 中每个向量 $\boldsymbol{\xi}$ 确定对偶空间 V^* 上的一个函数 $F_{\boldsymbol{\xi}}$, 它由 $F_{\boldsymbol{\xi}}(f)=(\boldsymbol{\xi},f)$ 来定义. 那么 (56) 表明 $F_{\boldsymbol{\xi}}$ 是线性函数.

定理 26 *任意有限维向量空间 V, 在下述对应下与它的二次共轭空间 $(V^*)^*$ 同构, 这个对应把每个向量 $\boldsymbol{\xi}\in V$ 映射到由 $F_{\boldsymbol{\xi}}(f)=\boldsymbol{\xi}f$ 定义的函数 $F_{\boldsymbol{\xi}}$ 上.*

证明 根据 (55) 式, 对应 $\tau:\boldsymbol{\xi}\longmapsto F_{\boldsymbol{\xi}}$ 保持向量加法和数乘运算. 我们现在证明 τ 是一一的, 因而是一个同构. 如果 $\boldsymbol{\xi}\neq\boldsymbol{\eta}$, 那么 $\boldsymbol{\zeta}=\boldsymbol{\xi}-\boldsymbol{\eta}\neq\mathbf{0}$, 于是 $\boldsymbol{\zeta}$ 是 V 的基的一部分. 因此, 根据定理 24, 在 V^* 中存在满足 $\boldsymbol{\zeta}f_0=1\neq0$ 的一个线性函数 f_0, 使得

$$F_{\boldsymbol{\xi}}(f_0)=F_{\boldsymbol{\eta}}(f_0)+F_{\boldsymbol{\zeta}}(f_0)=F_{\boldsymbol{\eta}}(f_0)+1\neq F_{\boldsymbol{\eta}}(f_0).$$

这就证明了 τ 是一一的, 因此它是 V 到 $(V^*)^*$ 的同构. 可是由定理 25 的推论 2, V 和 $(V^*)^*$ 的维数相同, 因此 τ 是映上的. 证毕

这个同构 $\boldsymbol{\xi}\longmapsto F_{\boldsymbol{\xi}}$, 同由推论 2 蕴含的 V 和 V^* 之间的同构不一样, 它是"自然的", 因为它的定义不依赖于 V 的基的选择.

设 S 是 V 的任意子空间, S' 是 V^* 中所有满足 $(\boldsymbol{\sigma},f)=0$(对每个 $\boldsymbol{\sigma}\in S$) 的那些线性函数 f 所组成的集合, 我们把集合 S' 同子空间 S 联系起来. 称 S' 是 S 的*零化子*. 显然它是 V^* 的子空间, 这因为由 $(\boldsymbol{\sigma},f)=0$ 和 $(\boldsymbol{\sigma},g)=0$ 可推出

$(\sigma, fc+gd)=0$. V 的子空间和它们在 V^* 中的零化子之间的对应 $S\to S'$ 具有性质

$$\text{若 } S\subset T \quad \text{可推出} \quad S'\supset T' \tag{57}$$

(包含关系是反的), 这因为如果 $f\in T'$, 那么对 T 中每个 $\boldsymbol{\sigma}$ 有 $(\boldsymbol{\sigma}, f)=0$, 因此对每个 $\boldsymbol{\sigma}\in S\subset T$ 也有 $(\boldsymbol{\sigma}, f)=0$. 仅有 $\mathbf{0}$ 一个元素组成的子空间的零化子是整个对偶空间 V^*, V 的零化子是 V^* 中仅有零函数组成的子空间.

由对偶性, 设 R 是共轭空间 V^* 的子空间, R' 是 V 中由所有满足 $(\boldsymbol{\xi}, f)=0$(对每个 $f\in R$) 的 $\boldsymbol{\xi}$ 组成的子空间, 则每个 R 确定一个 R' 作为它的零化子.

定理 27 如果 S 是 n 维向量空间 V 中的 k 维子空间, 那么把 S 零化的所有线性函数 f 组成的集合 S' 是 V^* 的 $n-k$ 维子空间.

证明 选取 S 的一组基 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_k$, 并根据定理 6 把它扩张成 V 的一组基 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n$. 在 V^* 的对偶基 $f_1,\cdots,f_n$ 中, 函数 $f_1c_1+\cdots+f_nc_n$ 对于 S 的所有向量都为零当且仅当它对于每个 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_k$ 为零, 也就是说, 当且仅当 $c_1=\cdots=c_k=0$. 这恰好意味着 $n-k$ 个函数 $f_{k+1},\cdots,f_k$ 构成 S 的零化子 S' 的一组基.

定理 27 恰是关于齐次线性方程组线性无关解的个数的定理 13 的重述.

子空间到它的零化子之间的对应 $S\longmapsto S'$ 引出 n 维射影几何的对偶原理, 在这方面, 下面的性质是基本的.

定理 28 对应 $S\longmapsto S'$ 满足

$$(S')'=S,\ (S+T)'=S'\cap T',\ (S\cap T)'=S'+T'. \tag{58}$$

证明 因为对 S 中一切 $\boldsymbol{\xi}$ 和 S' 中一切 $\boldsymbol{f}$, 有 $(\boldsymbol{\xi}, \boldsymbol{f})=0$, 所以 S 中每个 $\boldsymbol{\xi}$ 零化 S' 中每个向量 $\boldsymbol{f}$, 因此 $\boldsymbol{\xi}\in(S')'$, 于是 $(S')'\supset S$. 但是根据定理 27, $(S')'$ 的维数等于 $n-(n-k)=k=d[S]$; 因此 $(S')'\supset S$ 是不可能的, 必有 $(S')'=S$.

这个关系表明, 子空间到它的零化子的对应 $S\longmapsto S'$, 当作用两次时就是恒等对应; 因此这个对应有逆并且是一一映上. 因为根据 (57) 它也是相反的包含关系, 所以我们推出, 它把包含 S 和 T 的最小子空间 $S+T$ 映射到包含在 S' 和 T' 中的最大子空间 $S'\cap T'$, 并由对偶性得到 $(S\cap T)'=S'+T'$.

推论 1 设 $L(V)$ 是域上有限维向量空间 V 的所有子空间组成的集合. 存在一个 $L(V)$ 到自身的一一对应, 这个对应使包含关系相反并且满足 (58).

证明 设 V 中选取任意一组固定基 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n$. 对 V 的任意子空间 S, 设 S' 是所有满足下面条件的向量 $\boldsymbol{\eta}=y_1\boldsymbol{\beta}_1+\cdots+y_n\boldsymbol{\beta}_n$ 组成的集合:

$$x_1y_1+\cdots+x_ny_n=0,\ \text{对}S\text{中一切}\quad \boldsymbol{\xi}=(x_1\boldsymbol{\beta}_1+\cdots+x_n\boldsymbol{\beta}_n). \tag{59}$$

我们重复一下证明定理 27 和 (58) 式时用过的论证, 就可以得出所需要的结果.

注 1 在有限维欧几里得向量空间 E 的情形，存在一个从 E 到它的对偶空间 E^* 的自然同构，它可以通过内蕴内积 (ξ,η) 来定义. 对每个向量 $\boldsymbol{\eta}\in E$，公式 $\boldsymbol{\xi}f_{\boldsymbol{\eta}}=(\boldsymbol{\xi},\boldsymbol{\eta})$ 定义了一个 E 上的函数 $f_{\boldsymbol{\eta}}$，因为 $(\boldsymbol{\xi},\boldsymbol{\eta})$ 是双线性的，所以 $f_{\boldsymbol{\eta}}$ 是线性的. 容易证明，对应 $\boldsymbol{\eta}\longmapsto f_{\boldsymbol{\eta}}$ 是 E 到 E^* 上的同构.

注 2 对于无穷维向量空间 V，V 到 V^* 的同构一般来说并不存在. 例如，设 V 是所有序列 $\xi=(x_1,\cdots,x_n,\cdots)$ 组成的向量空间，其中各分量 $x_n\in F$，并且只有有限多个非零元素，加法和乘法是逐项进行的. V 上任意线性函数仍然可以表示成形式 $\boldsymbol{\xi}f=\Sigma x_ic_i$，其中的系数是任意无穷序列 $\gamma=(c_1,c_2,\cdots,c_n,\cdots)$. 因此对偶空间 V^* 是由所有这样的无穷序列组成. 空间 V 和 V^* 不同构. 例如，我们借助于更新的概念，如果 F 是可数域，那么 V 是可数的，但 V^* 却是不可数.

习 题

1. 完成定理 25 的证明.
2. 设 $f_1,\cdots,f_n$ 是 n 维向量空间 V 上 n 个线性无关的线性函数，$c_1,\cdots,c_n$ 是已知常数. 证明：V 中存在一个且只存在一个满足 $\boldsymbol{\xi}f_i=c_i(i=1,\cdots,n)$ 的向量 $\boldsymbol{\xi}$. 利用非齐次线性方程组加以解释.
3. (a) 完成注 1 的证明.

   (b) 指出注 1 与定理 25 的推论 1 的联系.
4. 在 $\mathbf{C}^4$ 中定义 $(\boldsymbol{\xi},\boldsymbol{\eta})=x_1y_2-y_1x_2+x_3y_4-y_3x_4$，对每个子空间 S，定义 S' 为所有满足 $(\boldsymbol{\xi},\boldsymbol{\eta})=0$(对一切 $\boldsymbol{\xi}\in S$) 的向量 $\boldsymbol{\eta}$ 组成的集合. 证明：(57) 和 (58) 成立. 并证明：如果 S 是一维空间，那么 $S\subset S'$.

第 8 章　矩 阵 代 数

8.1　线性变换与矩阵

有很多方法可以把一个平面线性地映射到自身, 也就是说, 使得向量的任意线性组合被映射到变换了的向量的同一线性组合. 用符号表示这就是

$$(c\boldsymbol{\xi} + d\boldsymbol{\eta})T = c(\boldsymbol{\xi}T) + d(\boldsymbol{\eta}T). \tag{1}$$

与此等价的说法是, T 按下述意义保持加法与数乘:

$$(\boldsymbol{\xi} + \boldsymbol{\eta})T = \boldsymbol{\xi}T + \boldsymbol{\eta}T, \quad (c\boldsymbol{\xi})T = c(\boldsymbol{\xi}T). \tag{2}$$

例如, 考虑平面围绕原点转过 θ 角的 (反时针) 刚体旋转 R_θ. 在几何上, 显然 R_θ 把以 $\boldsymbol{\xi}$ 和 $\boldsymbol{\eta}$ 为边的平行四边形的对角线 $\boldsymbol{\xi} + \boldsymbol{\eta}$ 变换到以 $\boldsymbol{\xi}R_\theta$ 和 $\boldsymbol{\eta}R_\theta$ 为边的平行四边形的对角线 $\boldsymbol{\xi}R_\theta + \boldsymbol{\eta}R_\theta$. 用图 8-1 加以说明, 在图中 $\theta = 135°$, 并可看出 $(\boldsymbol{\xi}+\boldsymbol{\eta})R_\theta = \boldsymbol{\xi}R_\theta + \boldsymbol{\eta}R_\theta$. 还有, 如果 c 是任意实标量, 则 $\boldsymbol{\xi}$ 的倍数 $c\boldsymbol{\xi}$ 旋转到 $c(\boldsymbol{\xi}R_\theta)$, 于是 $(c\boldsymbol{\xi})R_\theta = c(\boldsymbol{\xi}R_\theta)$. 因此平面的任意刚体旋转都是线性的, 此外, 对于空间围绕任一轴的旋转, 可做同样的考虑.

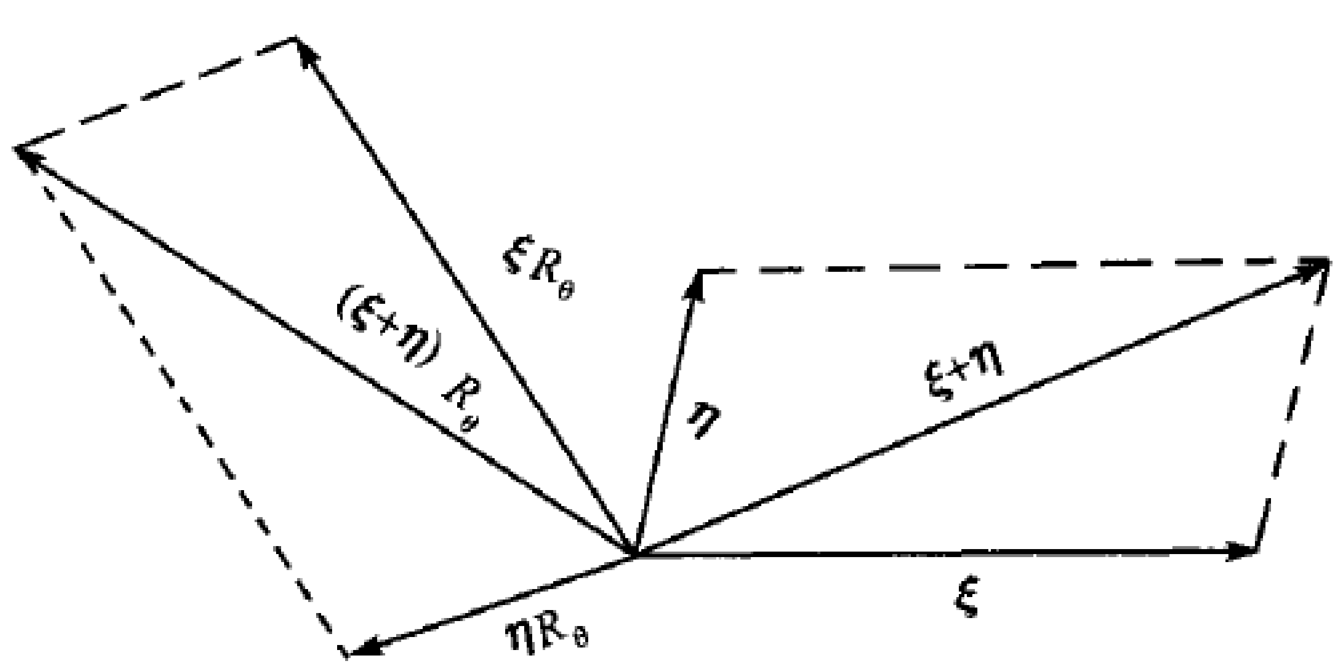


图　8-1

再有, 考虑平面离开原点的简单扩展 D_k, 在这个扩展下, 平面上每个点沿径向移动到一个位置, 这个位置到原点的距离是原距离的 k 倍, 用符号表示, 我们有

$$\boldsymbol{\xi}D_k = k\boldsymbol{\xi}, \qquad \text{对一切}\boldsymbol{\xi}. \tag{3}$$

这个变换又把平行四边形变为平行四边形, 因此向量和变为向量和, 所以 $(\boldsymbol{\xi}+\boldsymbol{\eta})D_k = \boldsymbol{\xi}D_k + \boldsymbol{\eta}D_k$. 此外, $(c\boldsymbol{\xi})D_k = kc\boldsymbol{\xi} = ck\boldsymbol{\xi} = c(\boldsymbol{\xi}D_k)$; 因此 D_k 是线性的. 注意, 如果

$0 < k < 1$, 那么公式 (3) 定义了一个朝向原点的简单压缩; 如果 $k = -1$, 那么公式 (3) 定义了一个关于原点的反射 (通过 180° 的旋转), 所以这些变换也都是线性的.

在任意有限维向量空间 F^n 中存在类似的变换, 比如, 设 T 是 $\mathbf{R}^3$ 的变换, 它把每个向量 $\boldsymbol{\xi} = (x_1, x_2, x_3)$ 变到向量 $\boldsymbol{\eta} = (y_1, y_2, y_3)$, $\boldsymbol{\eta}$ 的坐标由$\boldsymbol{\xi}$的坐标 x_1, x_2, x_3 的齐次线性函数给出:

$$\begin{aligned} y_1 &= a_1x_1 + b_1x_2 + c_1x_3, \\ y_2 &= a_2x_1 + b_2x_2 + c_2x_3, \\ y_3 &= a_3x_1 + b_3x_2 + c_3x_3. \end{aligned} \tag{4}$$

显然, 如果所有 x_i 都乘上同一个常数 d, 那么 (4) 中的所有 y_i 也同样乘上常数 d, 所以 $(d\boldsymbol{\xi})T = d\boldsymbol{\eta} = d(\boldsymbol{\xi}T)$. 同样, 向量 $\boldsymbol{\xi}$ 与 $\boldsymbol{\xi}' = (x_1', x_2', x_3')$ 的和 $\boldsymbol{\xi} + \boldsymbol{\xi}' = (x_1 + x_1', x_2 + x_2', x_3 + x_3')$ 的变换式 ξ, 可以由 (4) 计算出它的坐标为

$$\begin{aligned} z_j &= a_j(x_1 + x_1') + b_j(x_2 + x_2') + c_j(x_3 + x_3') \\ &= (a_jx_1 + b_jx_2 + c_jx_3) + (a_jx_1' + b_jx_2' + c_jx_3'), \end{aligned}$$

其中 $j = 1, 2, 3$. 这些 z_j 恰好等于 $y_j + y_j'$, 这里 y_j 由 (4) 式给出, 而 y_j' 相应于 (4) 的表示. 这就是说 $(\boldsymbol{\xi} + \boldsymbol{\xi}')T = \boldsymbol{\xi}T + \boldsymbol{\xi}'T$.

反过来, 在 $\mathbf{R}^3$ 上任意到自身的线性变换具有形式 (4). 为得到这个结论, 分别用

$$\boldsymbol{\alpha} = (a_1, a_2, a_3), \quad \boldsymbol{\beta} = (b_1, b_2, b_3), \quad \boldsymbol{\gamma} = (c_1, c_2, c_3)$$

表示单位向量

$$\boldsymbol{\varepsilon}_1 = (1, 0, 0), \quad \boldsymbol{\varepsilon}_2 = (0, 1, 0), \quad \boldsymbol{\varepsilon}_3 = (0, 0, 1)$$

的变换式, 那么变换 T 一定把 $\mathbf{R}^3$ 中每个向量 $\boldsymbol{\xi} = (x_1, x_2, x_3)$ 变到

$$\begin{aligned} \boldsymbol{\eta} = \boldsymbol{\xi}T &= (x_1\boldsymbol{\varepsilon}_1 + x_2\boldsymbol{\varepsilon}_2 + x_3\boldsymbol{\varepsilon}_3)T \\ &= x_1(\boldsymbol{\varepsilon}_1T) + x_2(\boldsymbol{\varepsilon}_2T) + x_3(\boldsymbol{\varepsilon}_3T) \\ &= x_1\boldsymbol{\alpha} + x_2\boldsymbol{\beta} + x_3\boldsymbol{\gamma} \\ &= (x_1a_1 + x_2b_1 + x_3c_1, x_1a_2 + x_2b_2 + x_3c_2, x_1a_3 + x_2b_3 + x_3c_3). \end{aligned}$$

因此, 如果 T 是线性的, 那么它就具有形式 (4).

前面的构造明显地给出 (4) 的系数, 比如, 考虑围绕原点转过 θ 角的反时针旋转 R_θ. 根据正弦函数和余弦函数的定义, 我们得到, 单位向量 $\boldsymbol{\varepsilon}_1 = (1, 0)$ 旋转到 $(\cos\theta, \sin\theta)$, 而单位向量 $\boldsymbol{\varepsilon}_2 = (0, 1)$ 旋转到

$$\left(\cos\left(\theta + \frac{\pi}{2}\right), \quad \sin\left(\theta + \frac{\pi}{2}\right)\right) = (-\sin\theta, \cos\theta).$$

这样, 在 (4) 中我们有 $a=\cos\theta, b=\sin\theta,\ a^*=-\sin\theta,\ b^*=\cos\theta$, 所以 R_θ 的方程是

$$R_\theta: x'=x\cos\theta-y\sin\theta,\quad y'=x\sin\theta+y\cos\theta. \tag{5}$$

同样, 关于通过原点并与 x 轴交成 α 角的直线的反射F_α, 它把极坐标是 (r,θ) 的点变换到极坐标是 $(r,2\alpha-\theta)$ 的点. 因此, 变换 F_α 的效果可用

$$F_\alpha: x'=x\cos 2\alpha+y\sin 2\alpha, y'=x\sin 2\alpha-y\cos 2\alpha \tag{5'}$$

表示.

关于线性的概念还可以更一般地应用到同一域上任意两个向量空间之间的变换.

定义　V 和 W 是同一域 F 上的向量空间, 变换 $T:V\to W$ 如果对 V 中一切向量 $\boldsymbol{\xi}$ 和 $\boldsymbol{\eta}$, 对 F 中一切标量 c 和 d, 有

$$(c\boldsymbol{\xi}+d\boldsymbol{\eta})T=c(\boldsymbol{\xi}T)+d(\boldsymbol{\eta}T),$$

那么称 T 为线性变换.

例如, 考虑变换

$$T_1:(x,y)\longmapsto(x+y,x-y,2x)=(x',y',z'), \tag{6}$$

它是由方程 $x'=x+y, y'=x-y, z'=2x$ 定义的. 这个变换把平面向量 (1, 0) 和 (0, 1) 分别变换到空间中的正交向量 (1, 1, 2) 和 (1, −1, 0), 并把平面线性地变换到空间的一个子集合.

用下面的原理处理有限维情况更为方便.

定理 1　如果 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_m$ 是向量空间 V 的任意一组基, $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 是向量空间 W 的任意 m 个向量, 那么存在唯一的线性变换 $T:V\to W$, 使得 $\boldsymbol{\beta}_1T=\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\beta}_mT=\boldsymbol{\alpha}_m$. 这个变换是通过

$$(x_1\boldsymbol{\beta}_1+\cdots+x_m\boldsymbol{\beta}_m)T=x_1\boldsymbol{\alpha}_1+\cdots+x_m\boldsymbol{\alpha}_m \tag{7}$$

来定义的.

例如, 在平面上, 设 $\boldsymbol{\beta}_1=(1,0),\ \ \boldsymbol{\beta}_2=(0,1),\ \ \boldsymbol{\alpha}_1=(1,0),\ \ \boldsymbol{\alpha}_2=(a,1)$. 那么定理 1 断言, 水平切变换

$$S_a:(x,y)\longmapsto(x+ay,y) \tag{8}$$

是线性变换, 并且是满足条件 $\boldsymbol{\beta}_1S_a=\boldsymbol{\alpha}_1,\ \ \boldsymbol{\beta}_2S_a=\boldsymbol{\alpha}_2$ 的唯一的线性变换, 几何上, 图上的每一点都沿 x 轴方向平行移动, 所通过的距离与这一点在 x 轴上面的高度成正比. 这个变换把其边平行于轴的矩形变换成平行四边形.(见图 8-2, 可以把它想象成一叠纸牌!)

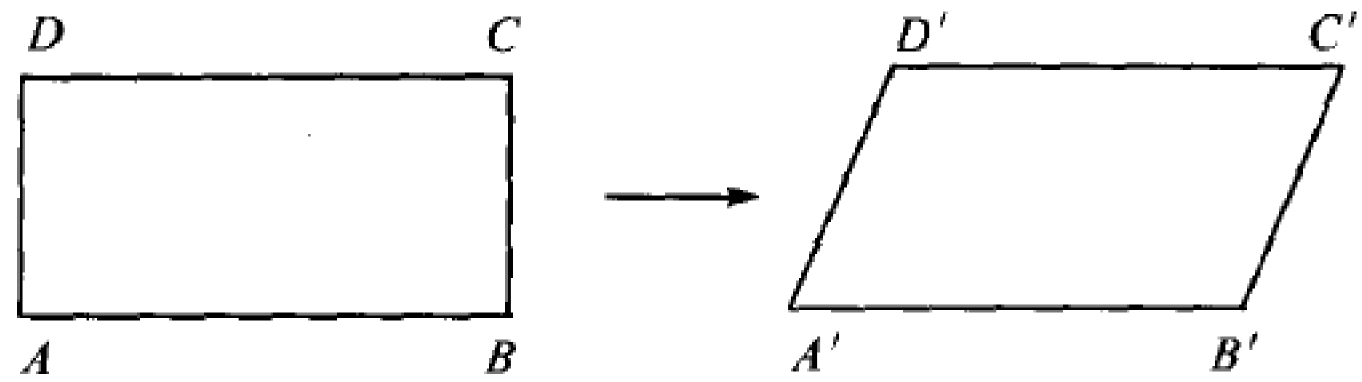


图 8-2

证明 如果 T 是线性变换, 并且 $\boldsymbol{\beta}_i T = \boldsymbol{\alpha}_i (i = 1, \cdots, m)$, 那么由定义 (1) 和归纳法得到公式 (7) 的明显形式. 另一方面, 因为 V 中每个向量可以唯一地表示为 $x_1\boldsymbol{\beta}_1 + \cdots + x_m\boldsymbol{\beta}_m$, 所以公式 (7) 定义了一个 V 到 W 的单值变换 T; 因此不可能存在另外的 V 到 W 的线性变换, 使 $\boldsymbol{\beta}_i T = \boldsymbol{\alpha}_i$. 为了证明 T 是线性的, 设 $\boldsymbol{\eta} = \sum y_i \boldsymbol{\beta}_i$ 是 V 中第二个向量, 那么,

$$\begin{aligned}(c\boldsymbol{\xi} + d\boldsymbol{\eta})T &= \left[\sum_{i=1}^{m} cx_i\boldsymbol{\beta}_i + \sum_{i=1}^{m} dy_i\boldsymbol{\beta}_i\right] T \\ &= \left[\sum_{i=1}^{m}(cx_i + dy_i)\boldsymbol{\beta}_i\right] T = \sum_{i=1}^{m}(cx_i + dy_i)\boldsymbol{\alpha}_i \\ &= c\sum_{i=1}^{m} x_i\boldsymbol{\alpha}_i + d\sum_{i=1}^{m} y_i\boldsymbol{\alpha}_i = c(\boldsymbol{\xi}T) + d(\boldsymbol{\eta}T).\end{aligned}$$

因此 T 是线性变换.

如果 $V = F^m, W = F^n$, 并设 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_m$ 是 V_m 的单位向量 $\boldsymbol{\varepsilon}_1 = (1, 0, \cdots, 0), \cdots, \boldsymbol{\varepsilon}_m = (0, 0, \cdots, 1)$, 我们得到定理 1 的一个非常重要的应用. 在这种情形, 我们可以给出每个 $\boldsymbol{\alpha}_i$ 的坐标表示,

$$\begin{aligned}\boldsymbol{\varepsilon}_1 T &= \boldsymbol{\alpha}_1 = (a_{11}, a_{12}, \cdots, a_{1n}), \\ \boldsymbol{\varepsilon}_2 T &= \boldsymbol{\alpha}_2 = (a_{21}, a_{22}, \cdots, a_{2n}), \\ &\ \vdots \qquad\qquad \vdots \\ \boldsymbol{\varepsilon}_m T &= \boldsymbol{\alpha}_m = (a_{m1}, a_{m2}, \cdots, a_{mn}).\end{aligned} \tag{9}$$

定理 1 指出, 刚好存在一个同公式 (9) 相联系的线性变换. 于是, 这个变换是由 $m \times n$ 矩阵 $\boldsymbol{A} = (a_{ij})$ 确定的, 矩阵 $\boldsymbol{A}$ 的第 i 行是一组坐标 $(a_{i1}, \cdots, a_{in})$, a_{ij} 是矩阵第 i 行第 j 列上的元素. 这样我们就证明了

定理 2 线性变换 $T : F^m \to F^n$ 和域 F 上的 $m \times n$ 矩阵 $\boldsymbol{A}$ 之间存在一个一一对应. 当给定变换 T, 相对应的矩阵 $\boldsymbol{A}$ 的第 i 行是由 $\boldsymbol{\varepsilon}_i T$ 的坐标组成; 当给定矩阵 $\boldsymbol{A} = (a_{ij})$, T 就是把 F^m 的每个单位向量 $\boldsymbol{\varepsilon}_i$ 变换到 $\boldsymbol{A}$ 的第 i 行 $(a_{i1}, \cdots, a_{in})$ 的唯一的线性变换.

我们用 $T_{\boldsymbol{A}}$ 表示 F^m 到 F^n 的按上述方式与 $\boldsymbol{A}$ 相对应的线性变换, 例如, 在平面上, (5) 式表示的旋转变换、(3) 式表示的相似变换和 (8) 式表示的切变换分别对应于下列矩阵

$$R_\theta \to \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \quad D_k \to \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}, \quad S_a \to \begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}.$$

(9) 式表示的一般变换 $T = T_{\boldsymbol{A}}$ 把 F^m 的任意已知向量 $\boldsymbol{\xi} = (x_1, \cdots, x_m) = x_1\boldsymbol{\varepsilon}_1 + \cdots + x_m\boldsymbol{\varepsilon}_m$ 变换到 $W = F^n$ 中的向量

$$\begin{aligned} \boldsymbol{\xi}T &= x_1\boldsymbol{\alpha}_1 + \cdots + x_m\boldsymbol{\alpha}_m \\ &= x_1(a_{11}, \cdots, a_{1n}) + \cdots + x_m(a_{m1}, \cdots, a_{mn}) \\ &= (x_1a_{11} + \cdots + x_ma_{m1}, \cdots, x_1a_{1n} + \cdots + x_ma_{mn}). \end{aligned}$$

因此, 如果 $(y_1, \cdots, y_n)$ 是变换了的向量 $\boldsymbol{\eta} = \boldsymbol{\xi}T$ 的坐标, 那么利用这些坐标通过一组齐次线性方程给出 T:

$$\begin{aligned} y_1 &= x_1a_{11} + x_2a_{21} + \cdots + x_ma_{m1} = \sum_i x_ia_{i1}, \\ y_2 &= x_1a_{12} + x_2a_{22} + \cdots + x_ma_{m2} = \sum_i x_ia_{i2}, \\ &\vdots \qquad\qquad \vdots \qquad\qquad\qquad\qquad \vdots \\ y_n &= x_1a_{1n} + x_2a_{2n} + \cdots + x_ma_{mn} = \sum_i x_ia_{in}. \end{aligned} \tag{10}$$

因此我们有

推论 从 F^m 到 F^n 的任意线性变换 T 可以用形为 (10) 的一组齐次线性方程来描述. 特别是, 每个 T 确定一个 $m \times n$ 矩阵 $\boldsymbol{A} = (a_{ij})$, 使得 T 把坐标为 $x_1, \cdots, x_n$ 的向量 $\boldsymbol{\xi}$ 变换到坐标为 $y_1, \cdots, y_n$ 的向量 $\boldsymbol{\eta} = \boldsymbol{\xi}T$, 其中 $y_1, \cdots, y_n$ 由 (10) 给出. 反过来, 每个 $m \times n$ 矩阵 $\boldsymbol{A}$ 通过方程 (10) 确定一个线性变换 $T = T_{\boldsymbol{A}} : F^m \to F^n$.

注意, (10) 的系数长方阵列并不是 (9) 中出现的矩阵 $\boldsymbol{A}$, 它是 (9) 式中行与列对换后的矩阵.(10) 的系数构成的 $n \times m$ 矩阵. 它是从 $m \times n$ 矩阵 $\boldsymbol{A}$ 通过行与列对换而得到的, 称为 $\boldsymbol{A}$ 的**转置**. 并用 $\boldsymbol{A}^{\mathrm{T}}$ 来表示. 如果 $\boldsymbol{A} = (a_{ij})$ 的第 i 行第 j 列元素为 a_{ij}, 那么矩阵 $\boldsymbol{A}$ 的转置 $\boldsymbol{B} = \boldsymbol{A}^{\mathrm{T}}$ 是通过关系

$$b_{ij} = a_{ij}^{\mathrm{T}} = a_{ji}(i = 1, \cdots, n; j = 1, \cdots, m) \tag{11}$$

形式地定义. 按照这种记号, 把 (10) 换成更熟悉的形式

$$\begin{aligned} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1m}x_m &= y_1, \\ b_{21}x_1 + b_{22}x_2 + \cdots + b_{2m}x_m &= y_2, \end{aligned}$$

$$\vdots \tag{11'}$$
$$b_{n1}x_1 + b_{n2}x_2 + \cdots + b_{nm}x_m = y_n.$$

上述线性变换公式涉及全体 m- 数组的空间 F^m 和全体 n- 数组的空间 F^n. 更一般地, 如果 V 和 W 是 F 上任意两个有限维向量空间, V 是 m 维的, W 是 n 维的, 当我们选取 V 的基 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_m$ 和 W 的基 $\gamma_1, \cdots, \gamma_n$ 之后, 那么任意线性变换 $T: V \to W$ 可以用矩阵 $\boldsymbol{A}$ 表示. 由于 T 是由像 $\boldsymbol{\beta}_i T = \sum_j a_{ij}\gamma_j$ 确定的, 所以我们说 T 是由对于已知基的这些系数组成的 $m \times n$ 矩阵 $\boldsymbol{A} = (a_{ij})$ 来表示. 这相当于, 在同构 $\Sigma x_i \boldsymbol{\beta}_i \longmapsto (x_1, \cdots, x_m), \Sigma y_j \gamma_j \longmapsto (y_1, \cdots, y_n)$ 之下, 用 m- 数组空间和 n- 数组空间分别代替空间 V 和空间 W.

习 题

1. 描述下列各线性变换的几何意义:
   (a) $y' = x, x' = y$; (b) $y' = x, x' = x$;
   (c) $y' = x, x' = 0$; (d) $y' = ky, x' = kx + kay$;
   (e) $y' = by, x' = cx$.
2. 考虑平面到自身的变换, 这些变换是按下面叙述的方式把每个点 P 变换到与 P 有关的点 P'. 确定什么时候变换是线性的, 并求出它的变换方程.
   (a) P' 在 P 的右方两个单位, 在 P 的上方一个单位 (平移).
   (b) P' 是 P 在过原点斜率为 $\frac{1}{2}$ 的直线上的投影.
   (c) P' 在连结 P 到原点的半直线 OP 上, P' 到 O 的距离满足 $\overline{OP'} = \dfrac{4}{\overline{OP}}$.
   (d) 把 P 围绕原点旋转 $30°$ 角, 接着再平行于 y 轴做切变换, 最后得到 P'.
   (e) P' 是 P 关于直线 $x = 3$ 的反射.
3. 求一矩阵, 表示顶点为 (1, 0) 和 $\left(-\dfrac{1}{2}, \pm\dfrac{\sqrt{3}}{2}\right)$ 的等边三角形的对称.
4. 描述下列空间线性变换的几何意义.
   (a) $x' = ax, y' = by, z' = cz$; (b) $x' = 0, y' = 3y, z' = 3z$;
   (c) $x' = x + 2y + 5z, y' = y, z' = z$; (d) $x' = x - y, y' = x + y, z' = 4z$.
5. 正文中 (6) 式的变换矩阵是什么?
6. 求出下列所描述的线性变换的矩阵:
   (a) $(1,1) \longmapsto (0,1),\ (-1,1) \longmapsto (3,2)$;
   (b) $(1,0) \longmapsto (4,0),\ (0,1) \longmapsto (-1,2)$;
   (c) $(2,3) \longmapsto (1,0),\ (3,2) \longmapsto (1,-1)$;
   (d) $(1,0,0) \longmapsto (1,2,1),\ (0,1,0) \longmapsto (3,1,1),\ (0,0,1) \longmapsto (0,0,3)$.

7. V 的子空间 S 在线性变换 T 之下的像, 指的是对 S 中一切 ξ 所有向量 ξT 构成的集合 $(S)T$. 证明 $(S)T$ 是一个子空间.
8. 一个线性变换 T 把 (1, 1) 变到 (0, 1, 2), 把 (−1, 1) 变到 (2, 1, 0). 表示 T 的矩阵是什么?

8.2 矩阵加法

线性变换 (矩阵) 的代数包含三种运算: 两个线性变换 (矩阵) 的加法、线性变换与标量的乘法 (数乘) 及两个线性变换 (矩阵) 的乘法. 我们现在定义矩阵的"向量"运算, 即两个矩阵的加法, 及矩阵与标量的乘法.

两个 $m \times n$ 矩阵 $\boldsymbol{A} = (a_{ij})$ 与 $\boldsymbol{B} = (b_{ij})$ 的和 $\boldsymbol{A} + \boldsymbol{B}$ 是通过相应元素相加而得到的, 表示为

$$(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij}). \tag{12}$$

这个和遵循通常的交换律和结合律, 因为矩阵和每个元素遵循这两个定律. 在这个加法运算之下, 所有元素为零的 $m \times n$ 矩阵 $\boldsymbol{O}$ 作为零矩阵, 所以

$$\boldsymbol{O} + \boldsymbol{A} = \boldsymbol{A} + \boldsymbol{O} = \boldsymbol{A}, \quad \text{对一切} m \times n \text{矩阵} \boldsymbol{A}.$$

矩阵的加法逆可以通过对每个元素简单地乘以 −1 而得到. 于是, 在加法运算之下, 全体 $m \times n$ 矩阵构成阿贝耳群.

矩阵 $\boldsymbol{A}$ 与标量 c 的"数乘"积 $c\boldsymbol{A}$ 是通过对矩阵的每个元素乘以 c 而得到的. 我们可以验证普通向量定律:

$$\begin{aligned} 1 \cdot \boldsymbol{A} = \boldsymbol{A}, \qquad & c(d\boldsymbol{A}) = (cd)\boldsymbol{A}, \\ (c+d)\boldsymbol{A} = c\boldsymbol{A} + d\boldsymbol{A}, \qquad & c(\boldsymbol{A} + \boldsymbol{B}) = c\boldsymbol{A} + c\boldsymbol{B}. \end{aligned} \tag{13}$$

定理 3 *在加法和数乘运算之下, 域 F 上的所有 $m \times n$ 矩阵构成一个 F 上的向量空间.*

任意矩阵 (a_{ij}) 可以写成和 $\Sigma a_{ij}E_{ij}$, 其中 E_{ij} 是一个特殊矩阵, 它的第 i 行第 j 列元素为 1, 其余元素都为 0. 这些矩阵 E_{ij} 是线性无关的, 所以它们构成所有 $m \times n$ 矩阵的空间的一组基. 因此这个空间的维数是 mn.

存在相应的线性变换的代数. 设 T 和 U 是从向量空间 V 到向量空间 W 的任意两个线性变换, 我们可以通过

$$\boldsymbol{\xi}(T + U) = \boldsymbol{\xi}T + \boldsymbol{\xi}U, \quad \text{对 } V \text{ 中一切 } \boldsymbol{\xi} \tag{14}$$

来定义和 $T+U$. 类似地, "数乘" 积 cT 通过 $\boldsymbol{\xi}(cT)=c(\boldsymbol{\xi}T)$ 来定义. 根据定义 (1), 线性变换之和 $T+U$ 是线性变换, 这因为

$$\begin{aligned}(c\boldsymbol{\xi}+d\boldsymbol{\eta})(T+U)&=(c\boldsymbol{\xi}+d\boldsymbol{\eta})T+(c\boldsymbol{\xi}+d\boldsymbol{\eta})U\\&=c\boldsymbol{\xi}T+c\boldsymbol{\xi}U+d\boldsymbol{\eta}T+d\boldsymbol{\eta}U\\&=c\boldsymbol{\xi}(T+U)+d\boldsymbol{\eta}(T+U).\end{aligned}$$

"数乘" 积cT 也是线性变换.

当 $V=F^m, W=F^n$ 时, 由定义 (14) 推出 $\boldsymbol{\varepsilon}_i(T+U)=\boldsymbol{\varepsilon}_iT+\boldsymbol{\varepsilon}_iU$, 因此按定理 2 那样, 与线性变换 $T+U$ 对应的矩阵 $\boldsymbol{C}$ 是与 T 和 U 对应的两个矩阵的和. 因此 $\boldsymbol{\varepsilon}_i(cT)=c(\boldsymbol{\varepsilon}_iT)$, 所以刚刚定义的线性变换数乘运算对应于前面定义的 $m\times n$ 矩阵的数乘运算. 按照定理 2 下面引进的记号, 这就是

$$T_{\boldsymbol{A}+\boldsymbol{B}}=T_{\boldsymbol{A}}+T_{\boldsymbol{B}}\quad 和\quad T_{c\boldsymbol{A}}=cT_{\boldsymbol{A}}. \tag{15}$$

这种新定义具有本质的优越性, 就是说, 这些定义与 V 和 W 中所用的坐标系无关 (参看 7.8 节). 它们还可以应用到无穷维向量空间.

最后, 应当看出, 向量空间 V 到向量空间 W 的线性变换恰好是 V 到 W 的同态 (V 和 W 都看作是阿贝耳群), 它们同样保持数乘运算. 由于这个原因, 所有从 V 到 W 的线性变换构成的向量空间常常记作 $\mathrm{Hom}(V,W)$.

习　题

1. 对 8.1 节的矩阵 R_θ, D_k, S_a, 计算 $2R_\theta+D_k, 2S_a-3D_k$ 和 $R_\theta-S_a+5D_k$.
2. 证明: $(\boldsymbol{A}+\boldsymbol{B})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}^{\mathrm{T}}, (c\boldsymbol{A})^{\mathrm{T}}=c\boldsymbol{A}^{\mathrm{T}}$.
3. 证明法则 (13).
4. 不借助于矩阵直接证明: 所有线性变换 $T: V\to W$ 的集合, 在 (14) 式和它下面所定义的加法及数乘运算之下是一个向量空间.

8.3 矩阵乘法

两个线性变换 T 和 U 的最重要的结合是它们的乘积 TU(像 6.2 节那样, 先作用 T, 后作用 U). 这一节中, 我们只考虑向量空间 V 到自身的两个线性变换 T 和 U 的乘积. 那么 TU 可以定义为 V 到自身的线性变换, 满足

$$\boldsymbol{\xi}(TU)=(\boldsymbol{\xi}T)U,\qquad 对\ V\ 中每个向量\boldsymbol{\xi}.$$

例如, 如果按 (8) 式的切变换之后再作用一个相似变换 D_k: 把 (x',y') 变到 $x''=kx',\ \ y''=ky'$, 它们的综合效果是把 (x,y) 变到 $x''=kx+kay,\ \ y''=ky$. 这个乘积 S_aD_k 仍然是线性的.

定理 4　两个线性变换的乘积是线性变换.

证明　按照定义, 乘积 TU 把任意 $\boldsymbol{\xi}$ 映射到 $\boldsymbol{\xi}(TU)=(\boldsymbol{\xi}T)U$. 由于 T 和 U 都是线性的, 所以有

$$\begin{aligned}(c\boldsymbol{\xi}+d\boldsymbol{\eta})TU&=[c(\boldsymbol{\xi}T)+d(\boldsymbol{\eta}T)]U\\&=c(\boldsymbol{\xi}TU)+d(\boldsymbol{\eta}TU),\end{aligned}\tag{16}$$

这个公式就是说, TU 也满足线性变换的定义条件 (1).

这个结论意味着, 对于 T 和 U 的两个齐次线性方程组 (10) 可以结合起来产生对于 TU 的齐次线性方程组. 具体地说, 设对应于矩阵 A 的变换

$$\begin{aligned}x'&=xa_{11}+ya_{21},\\y'&=xa_{12}+ya_{22},\end{aligned}\qquad \boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix},\tag{17}$$

后面有平面的第二个线性变换, 它把 (x',y') 映射到 (x'',y''), 其中

$$\begin{aligned}x''&=x'b_{11}+y'b_{21},\\y''&=x'b_{12}+y'b_{22},\end{aligned}\qquad \boldsymbol{B}=\begin{pmatrix}b_{11}&b_{12}\\b_{21}&b_{22}\end{pmatrix}.\tag{18}$$

把 (17) 代入 (18), 得到结合后的变换为

$$\begin{aligned}x''&=(a_{11}b_{11}+a_{12}b_{21})x+(a_{21}b_{11}+a_{22}b_{21})y,\\y''&=(a_{11}b_{12}+a_{12}b_{22})x+(a_{21}b_{12}+a_{22}b_{22})y.\end{aligned}\tag{19}$$

这个乘积变换的系数矩阵是由原来矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 根据下面重要法则计算出来:

$$\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix}\cdot\begin{pmatrix}b_{11}&b_{12}\\b_{21}&b_{22}\end{pmatrix}=\begin{pmatrix}a_{11}b_{11}+a_{12}b_{21}&a_{11}b_{12}+a_{12}b_{22}\\a_{21}b_{11}+a_{22}b_{21}&a_{21}b_{12}+a_{22}b_{22}\end{pmatrix}.\tag{20}$$

这个运算结果的第一行第二列元素只包含 $\boldsymbol{A}$ 的第一行元素和 $\boldsymbol{B}$ 的第二列元素, 等等. 这种乘法法则是个省事的方法, 它避免了由变量代换 (像 (19) 式那样) 而带来的麻烦.

对于 $n\times n$ 矩阵, 类似的公式也成立, 这是因为定理 2 和定理 4 指出, 变换 $T,U:F^n\to F^n$ 的乘积一定产生一个相应的矩阵的乘积. 我们现在来计算对应于 $T_{\boldsymbol{A}}T_{\boldsymbol{B}}$ 的矩阵的乘积 $\boldsymbol{AB}$, 以便给出法则

$$T_{\boldsymbol{A}}T_{\boldsymbol{B}}=T_{\boldsymbol{AB}}.\tag{21}$$

由定理 2, $\varepsilon_i T_{\boldsymbol{A}} = \sum_i a_{ij}\varepsilon_j$, $\varepsilon_j T_{\boldsymbol{B}} = \sum_k b_{jk}\varepsilon_k$. 因此

$$\begin{aligned}\varepsilon_i(T_{\boldsymbol{A}}T_{\boldsymbol{B}}) = (\varepsilon_i T_{\boldsymbol{A}})T_{\boldsymbol{B}} &= \left(\sum_j a_{ij}\varepsilon_j\right)T_{\boldsymbol{B}} = \sum_j a_{ij}(\varepsilon_j T_{\boldsymbol{B}}) \\ &= \sum_j a_{ij}\left(\sum_k b_{jk}\varepsilon_k\right) = \sum_k c_{ik}\varepsilon_k,\end{aligned}$$

其中

$$c_{ik} = \sum_j a_{ij}b_{jk} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk}. \tag{22}$$

因此, 为使 (21) 成立, 矩阵乘积 $\boldsymbol{C} = \boldsymbol{AB}$ 必须由 (22) 式来定义. 我们采用这个定义.

定义 *$n \times n$ 矩阵 $\boldsymbol{A}$ 和 $n \times n$ 矩阵 $\boldsymbol{B}$ 的乘积 $\boldsymbol{AB}$ 定义为 $n \times n$ 矩阵 $\boldsymbol{C}$, 它的第 i 行第 k 列元素 c_{ik} 由 (22) 式给出.*

两个矩阵的乘积也可以用语言来描述: 乘积 $\boldsymbol{AB}$ 的第 i 行第 k 列元素 c_{ik} 是通过 $\boldsymbol{A}$ 的第 i 行与 $\boldsymbol{B}$ 的第 k 列 "相乘" 而得到. 行与列 "相乘" 的意思是把它们相应的元素相乘然后再把结果相加.

从矩阵乘法和变换乘法之间的对应关系 (12) 直接推出, 矩阵乘法满足结合律. 用符号表示就是

$$\boldsymbol{A}(\boldsymbol{BC}) = (\boldsymbol{AB})\boldsymbol{C}. \tag{23}$$

这因为等式两边的矩阵分别对应于变换 $T_{\boldsymbol{A}}(T_{\boldsymbol{B}}T_{\boldsymbol{C}})$ 和 $(T_{\boldsymbol{A}}T_{\boldsymbol{B}})T_{\boldsymbol{C}}$, 根据变换乘法的结合律 (6.2 节), 它们是相等的.

矩阵乘法不仅满足结合律, 而且还满足对于矩阵和的分配律, 这因为矩阵 $(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C}$ 的元素 d_{ik} 由 (22) 那样的公式给出:

$$d_{ik} = \sum_j (a_{ij} + b_{ij})c_{jk} = \sum_j a_{ij}c_{jk} + \sum_j b_{ij}c_{jk}.$$

这就把 d_{ik} 分成 $\boldsymbol{AC}$ 的元素 g_{ik} 和 $\boldsymbol{BC}$ 的元素 h_{ik} 之和, 于是就证明了下面两个分配律中的第一个:

$$(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC} + \boldsymbol{BC}, \qquad \boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}. \tag{24}$$

对于与 d 的 "数乘" 积, 我们也可以验证下面定律

$$(d\boldsymbol{A})\boldsymbol{B} = d(\boldsymbol{AB}) \qquad \text{和} \qquad \boldsymbol{A}(d\boldsymbol{B}) = d(\boldsymbol{AB}). \tag{25}$$

把定律 (24) 和 (25) 概括起来就是说, 矩阵乘法是双线性的, 这是因为这些定律的前一半公式结合起来得到 $(dA + d^*A^*)B = d(AB) + d^*(A^*B)$. 这恰恰表明, 用矩阵 B 右乘是所有 $n \times n$ 矩阵 X 组成的向量空间上的一个线性变换 $X \longmapsto XB$. (24) 和 (25) 的另一半公式断言, 用矩阵 A 左乘也是线性变换.

与 F^n 的恒等变换 T_I 对应的是 $n \times n$单位矩阵 I, 它的主对角线上 (从左上方到右下方) 的元素 $e_{ii} = 1$, 而其他元素都是零. 这是因为对所有 $i = 1, \cdots, n$ 有 $\varepsilon_i T_I = \varepsilon_i$. 因为 I 表示恒等变换, 所以它具有性质: 对每个 $n \times n$ 矩阵 A, 有 $IA = A = AI$.

我们可以把前面的结论概括如下:

定理 5 域 F 上的所有 $n \times n$ 矩阵组成的集合在乘法之下是封闭的, 该乘法满足结合律, 具有单位元素, 并且对于向量加法和数乘运算是双线性的.

然而, 乘法不满足交换律, 比如

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

提示: 对于 6.1 节中的正方形, 这些矩阵引导出什么样的几何变换?

不是所有非零矩阵都有乘法逆元素, 例如矩阵 $\begin{pmatrix} 1 & 0 \\ 3 & 0 \end{pmatrix}$, 它表示一个在 x 轴上的斜投影, 不能引导出一一变换, 因而不是映上的; 所以它没有左逆元素和右逆元素 (6.2 节的定理 1). 类似地, 消去律也不成立, 因为存在很多零因子, 比如

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 2 & 4 \\ -1 & -2 \end{pmatrix}\begin{pmatrix} 0 & 2 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

中等号左边的矩阵.

公式 (15) 和 (21) 断言下面的重要原理.

定理 6 F^n 的线性变换代数与 F 上所有 $n \times n$ 矩阵代数, 在定理 2 的对应 $T_A \leftrightarrow A$ 之下是同构的.

这就暗示了, 定理 5 中所断言的关于矩阵代数的一些定律实际上对任意向量空间的线性变换也是正确的. 这种推测是容易验证的, 并且当我们应用于适当的无穷维向量空间时, 就直接导出 "运算微积" 的某些形式.

例 1 设 V 是由实变量 x 的所有函数 $f(x)$ 组成, J 是变换或 "算子" $[f(x)]J = f(x+1)$. 如果 I 是恒等变换, 则算子 $\Delta = J - I$ 称为 "差分算子", 它把 $f(x)$ 变换到 $f(x+1) - f(x)$. 算子 J 和 Δ 都是线性的, 这因为 $[cf(x) + dg(x)]J =$

$c[f(x)]J+d[g(x)]J$. 虽然可以应用关于线性的定义，但是我们注意，在无穷维空间中并不能建立齐次线性方程组. 对固定的 $a(x)$, 运算 $f(x)\to a(x)f(x)$ 也是线性的.

例 2 微分算子 D 用于 C^∞ 空间, C^∞ 是由具有任意阶导数的所有函数 $f(x)$ 组成, D 把 $f(x)$ 映射到 $f'(x)$, 它是线性的. 泰勒定理用符号表示可以写成 $\mathrm{e}^D=J$.

例 3 对于两个变量的函数 $f(x,y)$, 存在相应的线性算子 J_x, $J_y, D_x, D_y, \Delta_x, \Delta_y$. 例如, $[f(x,y)]J_x=f(x+1,y), [(f(x,y)]D_x=f'_x(x,y)$.

习　题

1. 对矩阵

$$\boldsymbol{A}=\begin{pmatrix}1&1\\0&1\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}1&-1\\0&1\end{pmatrix},\quad \boldsymbol{C}=\begin{pmatrix}1&3\\2&1\end{pmatrix},\quad \boldsymbol{D}=\begin{pmatrix}4&0\\2&1\end{pmatrix}$$

计算下列各式:

(a) $\boldsymbol{AB}, \boldsymbol{BA}, \boldsymbol{A}^2+\boldsymbol{AB}-2\boldsymbol{B}$;

(b) $(\boldsymbol{A}+\boldsymbol{B}-\boldsymbol{I})(\boldsymbol{A}-\boldsymbol{B}+\boldsymbol{I})-(\boldsymbol{A}+2\boldsymbol{B})(\boldsymbol{B}-\boldsymbol{A})$;

(c) $\boldsymbol{DB}, \boldsymbol{AC}, \boldsymbol{AD}$;

(d) 对于乘积 $(\boldsymbol{AC})\boldsymbol{D}, \boldsymbol{A}(\boldsymbol{CD})$ 验证结合律.

2. 利用矩阵乘法列出下列各变换 (按照 8.1 节中的记号) 方程:

(a) D_kS_a,　(b) S_aD_k,　(c) $R_\theta S_a(\theta=45^\circ)$,

(d) $R_\theta S_aD_k(\theta=30^\circ)$,　(e) $D_kS_aD_k$.

3. 何时 $S_aD_k=D_kS_a$(按习题 2 的记号)?

4. 用 T_n 表示 8.1 节习题 $4(n)(n=a,b,c,d)$ 所描述的变换. (用矩阵) 计算下列乘积:

(a) T_bT_c,　(b) T_aT_c,　(c)$T_bT_aT_b$,　(d)T_dT_c,　(e) $T_cT_bT_d$.

5. 证明定律 (25) 和定律 (24) 的第二个公式.

6. (a) 展开 $(\boldsymbol{A}+\boldsymbol{B})^3$,　(b) 证明 $\boldsymbol{A}^3\boldsymbol{A}^2=\boldsymbol{A}^2\boldsymbol{A}^3$.

7. 直接从定义 (22) 证明矩阵乘法的结合律.

8. 考虑两个矩阵的新“乘积” $\boldsymbol{A}\times\boldsymbol{B}$, 把它定义为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的“行与行”相乘. 这个乘积满足结合律吗?

9. (a) 设

$$\boldsymbol{B}=\begin{pmatrix}1&2&1\\1&3&2\\1&4&6\end{pmatrix},\quad \boldsymbol{E}_1=\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix},$$

$$\boldsymbol{E}_2=\begin{pmatrix}1&0&k\\0&1&0\\0&0&1\end{pmatrix},\quad \boldsymbol{E}_3=\begin{pmatrix}a&0&0\\0&b&0\\0&0&c\end{pmatrix}.$$

计算乘积 $\boldsymbol{BE}_1, \boldsymbol{BE}_2, \boldsymbol{BE}_3, \boldsymbol{E}_2\boldsymbol{E}_3, \boldsymbol{E}_1\boldsymbol{E}_3$.

(b) 设 $\boldsymbol{A}$ 是任意 3×3 矩阵, $\boldsymbol{AE}_3$ 与 $\boldsymbol{A}$ 有什么关系?

(c) 描述一下用 $\boldsymbol{E}_1$ 和 $\boldsymbol{E}_2$ 右乘任意 3×3 矩阵所产生的效果.

10. 不用矩阵证明：对于 V 到它自身的任意线性变换 R,S,T, 有定律

$$R(S+T)=RS+RT,\quad (R+S)T=RT+ST,\quad S(cT)=c(ST).$$

*11. 证明：如果 R, S, T 是向量空间的任意变换 (线性的或者不是线性的), 那么 $R(S+T)=RS+RT$ 成立, 但是 $(R+S)T=RT+ST$ 一般来说并不成立, 除非 T 是线性的.

12. 求出所有同习题 9 中的矩阵 $\boldsymbol{E}_3(a,b,c$ 不同) 可交换的矩阵.

*13. 证明：同习题 1 的矩阵 $\boldsymbol{D}$ 可交换的每个矩阵可以表示成形式

$$a\boldsymbol{I}+b\boldsymbol{D}.$$

14. 设 $\boldsymbol{A}$ 是任意 $n\times n$ 矩阵, 证明：所有同 $\boldsymbol{A}$ 可交换的 $n\times n$ 矩阵组成的集合 $C(\boldsymbol{A})$, 在加法和乘法运算之下是封闭的.

*15. 证明：每个 $n\times n$ 矩阵 $\boldsymbol{A}$ 满足形为

$$\boldsymbol{A}^m+c_{m-1}\boldsymbol{A}^{m-1}+\cdots+c_1\boldsymbol{A}+c_0\boldsymbol{I}=0,\quad m\leqslant n^2$$

的方程.

*16. (a) 设 $\boldsymbol{A}=(a_{ij})$ 是 $n\times n$ 实数矩阵, M 是 $|a_{ij}|$ 中最大的一个. 证明：$\boldsymbol{A}^k$ 的元素是有界的, 其数值不超过 $n^{k-1}M^k$.

(b) 证明：级数 $\boldsymbol{I}+\boldsymbol{A}+\dfrac{\boldsymbol{A}^2}{2!}+\dfrac{\boldsymbol{A}^3}{3!}+\cdots$ 总是收敛的. (它可以用来定义矩阵 $\boldsymbol{A}$ 的指数函数 $\mathrm{e}^{\boldsymbol{A}}$.)

在习题 17 ~ 习题 21 中采用上面例 1 ~ 例 3 的记号.

17. (a) 证明 D 是线性的. (b) 指出为什么 $\mathrm{e}^D=\boldsymbol{J}$.

18. 证明：$D_xD_y=D_yD_x$.

*19. (a) 化简 $xD-Dx, x\Delta-\Delta x, x\Delta^2-\Delta^2x$.

(b) 化简 $x^iD^j-D^jx^i, x^i\Delta^j-\Delta^jx^i$.

*20. 定义拉普拉斯算子 ∇^2 为 $\nabla^2=D_x^2+D_y^2$. 求

$$x\nabla^2-\nabla^2x, y(\nabla^2)^2-(\nabla^2)^2y, \nabla^2(x^2+y^2)-(x^2+y^2)\nabla^2.$$

*21. 用二项定理展开 $\Delta^n=(\boldsymbol{J}-\boldsymbol{I})^n$.

8.4 对角矩阵 · 置换矩阵 · 三角形矩阵

一个方阵 $\boldsymbol{D}=(d_{ij})$ 称为*对角矩阵*当且仅当对于 $i\neq j$ 时 $d_{ij}=0$; 也就是当且仅当 $\boldsymbol{D}$ 的所有非零元素都在主对角线 (从左上方到右下方) 上. 两个对角矩阵相加或相乘, 仅仅是对角线上相应的元素相加或相乘 (为什么?). 如果 $\boldsymbol{D}$ 的所有对角线

元素 d_{ii} 都非零, 那么对角矩阵 $\boldsymbol{E} = (e_{ij})$(其中 $e_{ii} = d_{ii}^{-1}$) 是 $\boldsymbol{D}$ 的逆, 这个逆是在 $\boldsymbol{DE} = \boldsymbol{I} = \boldsymbol{ED}$ 的意义下. 那么我们可以证明

定理 7 所有 $n \times n$ 对角矩阵, 如果其对角线元素都是域 F 上的非零元素, 那么它们在乘法之下构成交换群.

一个方阵 $\boldsymbol{P}$, 如果它的每一行和每一列都只是某个元素为 1, 其余元素都为零, 那么称 $\boldsymbol{P}$ 为*置换矩阵*

3×3 置换矩阵共有六个, 它们是单位矩阵 $\boldsymbol{I}$ 和矩阵

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

因为矩阵的行是单位向量的变换, 所以矩阵 $\boldsymbol{P}$ 是置换矩阵当且仅当它所对应的 V_n 的线性变换 T_p 是单位向量 $\varepsilon_1, \cdots, \varepsilon_n$ 的一个置换. 因此全体 $n \times n$ 置换矩阵与 n 个符号的 $n!$ 种可能的置换 (6.9 节) 一一对应, 并且这个对应是一个同构.

定理 8 全体 $n \times n$ 置换矩阵在乘法之下构成一个群, 它与 n 个字母的对称群同构.

还有另外一些重要的矩阵类. 如果矩阵 $\boldsymbol{M}$ 的每行和每列都恰有一个非零元素, 则称 $\boldsymbol{M}$ 是*单项矩阵*; 任意这样的矩阵可以通过把置换矩阵中的 1 用任意非零元素来代替而得到的, 例如

$$\boldsymbol{M}_1 = \begin{pmatrix} 0 & 0 & 5 \\ -2 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}, \quad \boldsymbol{M}_2 = \begin{pmatrix} 0 & 7 & 0 \\ 0 & 0 & -3 \\ 4 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{M}_3 = \begin{pmatrix} 0 & 4 \\ -1 & 0 \end{pmatrix}. \tag{26}$$

如果一个方阵 $\boldsymbol{T} = (t_{ij})$ 的对角线下面的元素都是零, 也就是如果当 $i > j$ 时, $t_{ij} = 0$, 则称 $\boldsymbol{T}$ 为*三角形矩阵*. 如果矩阵 $\boldsymbol{S}$ 的主对角线上的元素和主对角线下面的元素都是零, 那么称 $\boldsymbol{S}$ 为*严格三角形矩阵*. 这两种类型的矩阵, 在 4×4 的情况下可以示意地表示成

$$\boldsymbol{T} = \begin{pmatrix} q & r & s & t \\ 0 & u & v & w \\ 0 & 0 & x & y \\ 0 & 0 & 0 & z \end{pmatrix}, \qquad \boldsymbol{S} = \begin{pmatrix} 0 & u & v & w \\ 0 & 0 & x & y \\ 0 & 0 & 0 & z \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{27}$$

这里字母表示任意元素. 数量矩阵是指可以写成 cI 的矩阵, 这里 I 是单位矩阵.

这种按照矩阵的非零元素规定矩阵类型的方案, 不是构造矩阵群的唯一方法. 任意线性变换群都可以用相应的矩阵群来表示. 例如, 正方形对称群是由线性变换组成. 选取原点在正方形的中心, x 轴平行于正方形的一条边. 如果表示运动 R, R', H 和 D 的方程是通过 x 和 y 写出 (见 6.1 节中的描述), 那么它们给出的变换具有下面矩阵形式

$$R \to \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad R' \to \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$H \to \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad D \to \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

这个群的其他 4 个元素可以类似地表示出来. 6.4 节给出的这个群的乘法表可以通过这里相应的矩阵相乘来计算 (试一试!). 换句话说, 正方形对称群与 8 个 2×2 矩阵组成的群同构.

前面的例子表明, 已知矩阵 $\boldsymbol{A}$ 可以有逆 $\boldsymbol{A}^{-1}$, 使得 $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$. 这样的矩阵称为*非奇异矩阵*或*可逆矩阵*; 我们将在 8.6 节中系统地研究它们.

习 题

1. 一个 $n\times n$ 矩阵 $\boldsymbol{A}$ 在它的右边乘上一个对角矩阵$\boldsymbol{D}$, 其效果是什么?
2. 如果 $\boldsymbol{D}$ 是对角矩阵, 并且对角线上的所有元素都不同, 那么什么样的矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{D}$ 可交换 (何时 $\boldsymbol{AD} = \boldsymbol{DA}$)?
3. 证明: 主对角线上的元素是 1 的 2×2 三角形矩阵表示一个切变换.
4. 明显地列出全体 3×3 置换矩阵与对称群之间的同构.
5. 设 S_i 是由第 i 个单位向量 $\boldsymbol{\varepsilon}_i$ 张成的 V_n 的一维子空间. 证明: 非奇异矩阵 $\boldsymbol{D}$ 是对角矩阵当且仅当它所对应的线性变换 $\boldsymbol{T_D}$ 把每个子空间 S_i 映射到自身.
6. 对于单项矩阵, 作类似于习题 5 的描述.
7. (a) 证明: 单项矩阵 $\boldsymbol{M}$ 可以唯一地表示成形式 $\boldsymbol{M} = \boldsymbol{DP}$, 这里 $\boldsymbol{D}$ 是非奇异对角矩阵, $\boldsymbol{P}$ 是置换矩阵. (提示: 运用习题 5 和习题 6.)

   (b) 把正文中的矩阵 $\boldsymbol{M}_1$ 和 $\boldsymbol{M}_2$ 写成形式 $\boldsymbol{DP}$ 和 $\boldsymbol{PD}$.

   *(c) 列出单项矩阵群到置换矩阵群上的同态映射.
8. 描述单项矩阵 $\boldsymbol{M}$ 的逆, 并求出 (26) 式中矩阵 $\boldsymbol{M}_1$ 和 $\boldsymbol{M}_2$ 的逆.
9. 设 $\boldsymbol{M}$ 是单项矩阵, $\boldsymbol{D}$ 是对角矩阵, 证明: $\boldsymbol{M}^{-1}\boldsymbol{DM}$ 是对角矩阵.
10. 设 $\boldsymbol{P}$ 是置换矩阵, $\boldsymbol{D}$ 是对角矩阵, 明显地描述变换 $\boldsymbol{P}^{-1}\boldsymbol{DP}$ 的形式.
11. 设 $\boldsymbol{P}$ 是置换矩阵, $\boldsymbol{PA}$ 的行与 $\boldsymbol{A}$ 的行之间有怎样的关系?
12. 如果矩阵 $\boldsymbol{A}$ 的某个幂是 $\boldsymbol{O}$, 则称 $\boldsymbol{A}$ 是幂零矩阵. 证明: 任意严格三角形矩阵是幂零矩阵. (提示: 试验 3×3 的情形.)

13. 把矩形对称群表示为矩阵群.
14. 对于以 $(\pm1,\pm1)$ 为顶点的正方形对称群, 计算出表示对称 H, D, V 的矩阵. 验证 $HD = DV$.
*15. 在习题 7 中证明: 公式 $\boldsymbol{M} = \boldsymbol{DP}$ 定义了一个群同态 $\boldsymbol{M} \to \boldsymbol{P}$. 求出它的核.

8.5 长方矩阵

迄今我们只考虑了 $n\times n$ 方阵的乘法, 现在我们讨论长方矩阵的乘法, 也就是 $m\times n$ 矩阵的乘法, 这里一般假定 $m \neq n$.

一个 $m\times n$ 矩阵 $\boldsymbol{A} = (a_{ij})$ 和一个 $n\times r$ 矩阵 $\boldsymbol{B} = (b_{jk})$, 它们具有相同的 n, 它们确定一个 $m\times r$ 矩阵 $\boldsymbol{C}$ 作为它们的乘积 $\boldsymbol{AB} = (c_{ik}), \boldsymbol{C}$ 的元素为

$$c_{ik} = \sum_j a_{ij}b_{jk},$$

这里对 j 从 1 到 n 求和. 这种 "行与列" 的乘积, 只有当 $\boldsymbol{A}$ 的每行长度恰好与 $\boldsymbol{B}$ 的每列长度一样时, 才能构成, 因此必须假定 $\boldsymbol{A}$ 的列数 n 等于 $\boldsymbol{B}$ 的行数 n. 比如, $m = 1, n = 2, r = 3$,

$$(x_1, x_2)\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = (x_1a_{11} + x_2a_{21}, x_1a_{12} + x_2a_{22}, x_1a_{13} + x_2a_{23}).$$

同上面公式 (21) 和公式 (22) 一样, 在定理 2 的意义下, 矩阵乘积 $\boldsymbol{AB}$ 对应于线性变换 $T_{\boldsymbol{A}} : F^m \to F^n$ 和 $T_{\boldsymbol{B}} : F^n \to F^r$ 的乘积 $T_{\boldsymbol{A}}T_{\boldsymbol{B}}$, 其中 $T_{\boldsymbol{A}}$ 和 $T_{\boldsymbol{B}}$ 分别与矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相对应. 这里我们像通常那样用

$$\boldsymbol{\xi}(TU) = (\boldsymbol{\xi}T)U, \qquad 对V中一切\boldsymbol{\xi} \tag{28}$$

来定义线性变换 $T : V \to W$ 与线性变换 $U : W \to X$ 的乘积.

对于方阵成立的代数定律, 对于长方矩阵也都成立, 只要假定这些长方矩阵具有合适的维数, 以保证所有的乘法都满足定义. 例如, $m\times m$ 单位矩阵 $\boldsymbol{I}_m$ 和 $n\times n$ 单位矩阵 $\boldsymbol{I}_n$ 满足

$$\boldsymbol{I}_m\boldsymbol{A} = \boldsymbol{A} = \boldsymbol{A}\boldsymbol{I}_n \qquad (当\boldsymbol{A}是\ m\times n\ 矩阵). \tag{29}$$

长方矩阵的乘法同 (24) 和 (25) 那样, 仍然是双线性的. 结合律是

$$\boldsymbol{A}(\boldsymbol{BC}) = (\boldsymbol{AB})\boldsymbol{C} \qquad (\boldsymbol{A}是m\times n矩阵, \boldsymbol{B}是n\times r矩阵, \boldsymbol{C}是\ r\times s\ 矩阵). \tag{30}$$

另外, 证明这个定律最好把长方矩阵解释为线性变换.

像 (11) 式那样, $m \times n$ 矩阵$\boldsymbol{A}$ 的转置是 $n \times m$ 矩阵 $\boldsymbol{A}^{\mathrm{T}}$, 它的元素 $a_{ij}^{\mathrm{T}} = a_{ji}(i = 1, \cdots, n; j = 1, \cdots, m)$. 转置矩阵 $\boldsymbol{A}^{\mathrm{T}}$ 的第 i 行是原矩阵 $\boldsymbol{A}$ 的第 i 列, 反之亦然. 我们把矩阵 $\boldsymbol{A}$ 按它的主对角线反射也可以得到转置矩阵 $\boldsymbol{A}^{\mathrm{T}}$. 为了计算乘积 $\boldsymbol{AB} = \boldsymbol{C}$ 的转置矩阵 $\boldsymbol{C}^{\mathrm{T}}$, 我们用公式

$$c_{ik}^{\mathrm{T}} = c_{ki} = \sum_j a_{kj} b_{ji} = \sum_j b_{ji} a_{kj} = \sum_j b_{ij}^{\mathrm{T}} a_{jk}^{\mathrm{T}}. \tag{31}$$

这个结果恰好是乘积 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$ 的 (i, k) 上的元素. (注意, 反序.) 这就证明了下面的第一个定律:

$$(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}, \quad (\boldsymbol{A} + \boldsymbol{B})^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}^{\mathrm{T}}, \quad (c\boldsymbol{A})^{\mathrm{T}} = c\boldsymbol{A}^{\mathrm{T}}. \tag{32}$$

因此, 对应 $\boldsymbol{A} \to \boldsymbol{A}^{\mathrm{T}}$ 保持了和并使乘积反序, 故有时也称为反自同构. 因为 $(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{A}$, 所以这个反自同构称为“对合”的反自同构.

全都使用长方矩阵表示有几个优点. 例如, F 上 n- 数组的空间 F^n 中的向量 $\boldsymbol{\xi}$ 可以看作正好是一行的 $1 \times n$ 矩阵 $\boldsymbol{X}$, 或称为“行矩阵”. 这允许我们把 (10) 式所定义的方程组 $y_j = \Sigma x_i a_{ij}$ 解释为行矩阵 $\boldsymbol{Y}$ 是行矩阵$\boldsymbol{X}$ 与矩阵 $\boldsymbol{A}$ 的乘积. 这样, 线性变换 $T_{\boldsymbol{A}} : F^m \to F^n$ 就可以缩写成形式

$$\boldsymbol{Y} = \boldsymbol{XA}, \quad \boldsymbol{X} \in F^m, \quad \boldsymbol{Y} \in F^n. \tag{33}$$

还有, “数乘” 积 $c\boldsymbol{X}$ 正好是 1×1 矩阵 $\boldsymbol{c}$ 与 $1 \times n$(行) 矩阵 $\boldsymbol{X}$ 的乘积.

列向量 注意, 虽然在方程 $\boldsymbol{XA} = \boldsymbol{Y}$ 中 $\boldsymbol{Y}$ 是一个行向量, 但它的元素在表达式 (10) 中却以单列的形式出现. 因此通常把矩阵方程 $\boldsymbol{XA} = \boldsymbol{Y}$ 改写成转置形式 $\boldsymbol{Y}^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}$, 这里 $\boldsymbol{Y}^{\mathrm{T}}$ 和 $\boldsymbol{X}^{\mathrm{T}}$ 都是列向量. 改变记号后, 结果得到 (11′) 形式的方程 $\boldsymbol{BX} = \boldsymbol{Y}$, 其中 $\boldsymbol{B} = \boldsymbol{A}^{\mathrm{T}}, \boldsymbol{X} = (x_1, \cdots, x_n)^{\mathrm{T}}, \boldsymbol{Y} = (y_1, \cdots, y_n)^{\mathrm{T}}, \boldsymbol{X}$ 和 $\boldsymbol{Y}$ 都是列向量.

在处理双线性型和二次型时, 行向量和列向量要一起使用. 例如, 两个向量的内积 $x_1y_1 + \cdots + x_ny_n$(7.9 节) 仅仅是行矩阵 $\boldsymbol{X}$ 和列矩阵 $\boldsymbol{Y}^{\mathrm{T}}$ 的乘积, 因此

$$(\boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{XY}^{\mathrm{T}}, \quad \boldsymbol{X} \text{ 和 } \boldsymbol{Y} \text{ 是行矩阵}. \tag{34}$$

矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的行与列的乘法实际上是 $\boldsymbol{A}$ 的第 i 行与 $\boldsymbol{B}$ 的第 k 列的矩阵乘法, 于是矩阵乘积的定义可写成

$$\boldsymbol{AB} = (c_{ik}), \quad \text{其中} c_{ik} = \boldsymbol{A}_i\boldsymbol{B}^{(k)} \tag{35}$$

这里我们使用了记号

$$\boldsymbol{A}_i = \boldsymbol{A} \text{的第 } i \text{ 行}, \quad \boldsymbol{B}^{(k)} = \boldsymbol{B} \text{ 的第 } k \text{ 列}. \tag{36}$$

乘积 $\boldsymbol{AB}$ 的整个第 i 行 $(c_{i1},\cdots,c_{in})$ 只用了 $\boldsymbol{A}$ 的第 i 行和 $\boldsymbol{B}$ 的各列, 因此它是 $\boldsymbol{A}_i$ 与整个 $\boldsymbol{B}$ 的乘积. 类似地, $\boldsymbol{AB}$ 的第 k 列只由 $\boldsymbol{B}$ 的第 k 列产生出来. 用 (36) 的记号, 这些运算法则就是

$$(\boldsymbol{AB})_i=\boldsymbol{A}_i\boldsymbol{B},\ \ (\boldsymbol{AB})^{(k)}=\boldsymbol{AB}^{(k)}. \tag{37}$$

通过把矩阵 $\boldsymbol{B}$ 写成以它的列为元素的行矩阵, 可以把第二个法则具体化, 也就是

$$\boldsymbol{A}(\boldsymbol{B}^{(1)}\ \ \boldsymbol{B}^{(2)}\ \ \cdots\ \ \boldsymbol{B}^{(r)})=(\boldsymbol{AB}^{(1)}\ \ \boldsymbol{AB}^{(2)}\ \ \cdots\ \ \boldsymbol{AB}^{(r)}). \tag{38}$$

这些列还可以分组, 构成较大的子矩阵. 比如, 6×5 矩阵 $\boldsymbol{B}$ 可以看作 6×2 矩阵 $\boldsymbol{D}_1=(\boldsymbol{B}^{(1)}\ \ \boldsymbol{B}^{(2)})$ 和 6×3 矩阵 $\boldsymbol{D}_2=(\boldsymbol{B}^{(3)}\ \boldsymbol{B}^{(4)}\ \boldsymbol{B}^{(5)})$ 并列构成整个 6×5 矩阵 $\boldsymbol{B}=(\boldsymbol{D}_1\ \ \ \boldsymbol{D}_2)$. 由 (38), 乘法法则变为

$$\boldsymbol{A}(\boldsymbol{D}_1\ \ \boldsymbol{D}_2)=(\boldsymbol{AD}_1\ \ \boldsymbol{AD}_2),\quad \boldsymbol{D}_1 \text{ 和 } \boldsymbol{D}_2 \text{ 是 } n \text{ 行的矩阵块}. \tag{39}$$

如果我们把 $n\times r$ 矩阵 $\boldsymbol{B}$ 分成 n 个行 $\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n$, 而 $\boldsymbol{Y}=(y_1,\cdots,y_n)$ 是行矩阵, 那么乘积 $\boldsymbol{YB}$ 是行矩阵

$$\begin{aligned}\boldsymbol{YB}&=(y_1b_{11}+\cdots+y_nb_{n1},\cdots,y_1b_{1r}+\cdots+y_nb_{nr})\\&=y_1(b_{11},\cdots,b_{1r})+\cdots+y_n(b_{n1},\cdots,b_{nr})\\&=y_1\boldsymbol{B}_1+\cdots+y_n\boldsymbol{B}_n.\end{aligned}$$

于是乘积 $\boldsymbol{YB}$ 由行 $\boldsymbol{Y}$ 与行 $\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n$ 组成的列相乘而构成. 例如, $\boldsymbol{AB}$ 的第 i 行是由行矩阵 $\boldsymbol{A}_i=(a_{i1},\cdots,a_{in})$ 与 $\boldsymbol{B}$ 的乘积来定义, 因此

$$(\boldsymbol{AB})_i=a_{i1}\boldsymbol{B}_1+\cdots+a_{in}\boldsymbol{B}_n,\quad i=1,\cdots,m. \tag{40}$$

于是 $\boldsymbol{AB}$ 的每一行是 $\boldsymbol{B}$ 的所有行的线性组合. 这些公式是矩阵通过分成 "块" 或子矩阵相乘的方法的特例. 概括描述这个方法的其他情形也是很方便的.

$$\left(\underbrace{\begin{matrix}a_{11}&\cdots&a_{1s}\\\vdots&&\vdots\\a_{m1}&\cdots&a_{ms}\end{matrix}}_{\boldsymbol{M}_1}\ \middle|\ \underbrace{\begin{matrix}a_{1,s+1}&\cdots&a_{1n}\\\vdots&&\vdots\\a_{m,s+1}&\cdots&a_{mn}\end{matrix}}_{\boldsymbol{M}_2}\right)\left(\begin{array}{ccc}b_{11}&\cdots&b_{1r}\\\vdots&&\vdots\\b_{s1}&\cdots&b_{sr}\\\hline b_{s+1,1}&\cdots&b_{s+1,r}\\\vdots&&\vdots\\b_{n1}&\cdots&b_{nr}\end{array}\right)\begin{matrix}\left.\right\}\boldsymbol{N}_1\\\\\left.\right\}\boldsymbol{N}_2\end{matrix}$$

设矩阵 $\boldsymbol{A}$ 的 n 列是由子矩阵 $\boldsymbol{M}_1$ 的 s 列和它后面子矩阵$\boldsymbol{M}_2$ 的 $n-s$ 列组成. 把矩阵 $\boldsymbol{B}$ 的行相应地分开, 使得 $s\times r$ 矩阵 $\boldsymbol{N}_1$ 在 $(n-s)\times r$ 矩阵 $\boldsymbol{N}_2$ 的顶上. 关

于 $\boldsymbol{AB}=\boldsymbol{C}$ 的乘积公式分成两个相应的部分

$$c_{ik}=(a_{i1}b_{1k}+\cdots+a_{is}b_{sk})+(a_{i,s+1}b_{s+1,k}+\cdots+a_{in}b_{nk}). \tag{41}$$

第一个括号中只用了 $\boldsymbol{A}$ 的第一块 $\boldsymbol{M}_1$ 的第 i 行和 $\boldsymbol{B}$ 的上边一块 $\boldsymbol{N}_1$ 的第 k 列, 因此第一个括号实际上是乘积块 $\boldsymbol{M}_1\boldsymbol{N}_1$ 的第 i 行第 k 列元素 d_{ik}. 同样, (41) 的第二个括号是乘积 $\boldsymbol{M}_2\boldsymbol{N}_2$ 的元素 d_{ik}^*. 因此 $c_{ik}=d_{ik}+d_{ik}^*$, 这样整个乘积 $\boldsymbol{AB}$ 是矩阵和 $\boldsymbol{M}_1\boldsymbol{N}_1+\boldsymbol{M}_2\boldsymbol{N}_2$. 这就是

$$(\boldsymbol{M}_1\quad \boldsymbol{M}_2)\begin{pmatrix}\boldsymbol{N}_1\\ \boldsymbol{N}_2\end{pmatrix}=\boldsymbol{M}_1\boldsymbol{N}_1+\boldsymbol{M}_2\boldsymbol{N}_2. \tag{42}$$

这个公式是分块矩阵的行与列的乘法公式, 这恰好同矩阵元素的行与列乘法一样. 如果把 $\boldsymbol{A}$ 的行任意分开, 同时把 $\boldsymbol{B}$ 的列也相应地分开, 那么类似的结果成立. 如果行和列都分开, 那么把公式 (42) 和法则 (39) 结合起来就得到分块矩阵乘法公式

$$\begin{aligned}&\begin{pmatrix}\boldsymbol{M}_{11} & \boldsymbol{M}_{12}\\ \boldsymbol{M}_{21} & \boldsymbol{M}_{22}\end{pmatrix}\begin{pmatrix}\boldsymbol{N}_{11} & \boldsymbol{N}_{12}\\ \boldsymbol{N}_{21} & \boldsymbol{N}_{22}\end{pmatrix}\\ =&\begin{pmatrix}\boldsymbol{M}_{11}\boldsymbol{N}_{11}+\boldsymbol{M}_{12}\boldsymbol{N}_{21} & \boldsymbol{M}_{11}\boldsymbol{N}_{12}+\boldsymbol{M}_{12}\boldsymbol{N}_{22}\\ \boldsymbol{M}_{21}\boldsymbol{N}_{11}+\boldsymbol{M}_{22}\boldsymbol{N}_{21} & \boldsymbol{M}_{21}\boldsymbol{N}_{12}+\boldsymbol{M}_{22}\boldsymbol{N}_{22}\end{pmatrix}.\end{aligned} \tag{43}$$

假定这里的划分满足: $\boldsymbol{M}_{11}$ 的列数等于 $\boldsymbol{N}_{11}$ 的行数. 运算法则 (43) 恰好同 8.3 节公式 (20) 关于 2×2 矩阵乘法法则一样, 只不过这里的元素 $\boldsymbol{M}_{ij}$ 和 $\boldsymbol{N}_{ij}$ 是子矩阵或矩阵块, 而不是标量. 因此我们得出结论: 在适当划分子块后, 分块矩阵乘法与普通矩阵乘法完全一样.

习 题

1. 设

$$\boldsymbol{A}=\begin{pmatrix}1 & 0 & 0\\ 0 & \frac{1}{2} & \frac{1}{2}\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}\mathrm{i} & 0 & -\mathrm{i}\\ 0 & 1 & 1+\mathrm{i}\end{pmatrix},\quad \boldsymbol{X}=(1,\ -1),\quad \boldsymbol{Y}=(\mathrm{i},\ 0).$$

   (a) 求 $\boldsymbol{XA},\boldsymbol{XB},\boldsymbol{YA},\boldsymbol{YB}$.

   (b) 求 $3\boldsymbol{A}-4\boldsymbol{B},\boldsymbol{A}+(1+\mathrm{i})\boldsymbol{B},[\boldsymbol{X}-(1+\mathrm{i})\boldsymbol{Y}](\mathrm{i}\boldsymbol{A}+5\boldsymbol{B})$.

   (c) 求 $\boldsymbol{BA}^{\mathrm{T}},\boldsymbol{AB}^{\mathrm{T}},\boldsymbol{XAB}^{\mathrm{T}}\boldsymbol{BA}^{\mathrm{T}}\boldsymbol{Y}^{\mathrm{T}}$.

2. 证明: 如果 $\boldsymbol{X}$ 是任意行向量, 那么 $\boldsymbol{XX}^{\mathrm{T}}$ 是 $\boldsymbol{X}$ 同它自身的内积, 而 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ 是以 $a_{ij}=x_ix_j$ 为元素的矩阵 $\boldsymbol{A}$.

3. 设

$$A=\begin{pmatrix}2&3&0&0\\5&2&0&0\\0&0&4&0\\0&0&0&2\end{pmatrix},\quad B=\begin{pmatrix}1&0&0&0\\0&1&0&0\\1&2&1&0\\3&4&0&1\end{pmatrix},\quad C=\begin{pmatrix}0&1\\1&0\\2&0\\0&2\end{pmatrix},$$

求 $\boldsymbol{AB},\boldsymbol{BA},\boldsymbol{AC}$ 和 $\boldsymbol{BC}$.

4. 设 $\boldsymbol{I}^*$ 是 $(r+n)\times n$ 矩阵, 它是由 $r\times n$ 零矩阵上面再放上一个 $n\times n$ 单位矩阵而构成. 任意一个 $n\times(r+n)$ 矩阵用 $\boldsymbol{I}^*$ 来乘其效果如何?

*5. 证明分块矩阵乘法法则 (43).

8.6 逆 矩 阵

有限维向量空间的线性变换分为两类: 或者是双射 (一一的和映上), 或者既不是单射也不是满射 (既不是一一映入也不是映上). 例如, 三维欧几里得空间到 (x,y) 平面上的斜投影 $(x,y,z)\longmapsto(x,y+z,0)$, 既不是单射也不是满射.

定义 *向量空间 V 到它自身的线性变换 T, 如果它是 V 到 V 上的双射, 则称 T 是非奇异的或可逆的. 否则, 称 T 是奇异的.*

非奇异的线性变换T 是从 V 到 V 上的双射, 它保持加法和"数乘"积两种代数运算, 因此它是向量空间 V 到它自身的同构. 所以 V 的非奇异线性变换可以称为 V 的自同构.

判断一个线性变换是奇异的还是非奇异的这一重要事实, 最直接的方法是使用第 7 章推导出的线性无关理论, 也就是, 利用向量空间 V 的一组固定基 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$, 让已知的线性变换 T 在这组基上进行运算, 最后判断无关性.

定理 9 *具有有限基 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 的向量空间 V 的线性变换 T 是非奇异的当且仅当向量 $\boldsymbol{\alpha}_1T,\cdots,\boldsymbol{\alpha}_nT$ 在 V 中线性无关. 如果 T 是非奇异的, 那么 T 有一个 (双边) 线性逆 T^{-1}, 满足 $TT^{-1}=T^{-1}T=I$*

证明 首先假定 T 是非奇异的. 如果 $\boldsymbol{\alpha}_1T,\cdots,\alpha_nT$ 之间存在一个线性关系 $\Sigma x_i(\boldsymbol{\alpha}_iT)=\mathbf{0}$, 那么

$$(x_1\boldsymbol{\alpha}_1+\cdots+x_n\boldsymbol{\alpha}_n)T=x_1(\boldsymbol{\alpha}_1T)+\cdots+x_n(\boldsymbol{\alpha}_nT)=\mathbf{0}.$$

因为 $\mathbf{0}T=\mathbf{0}$, 并且 T 是一一的, 这就推出 $x_1\boldsymbol{\alpha}_1+\cdots+x_n\boldsymbol{\alpha}_n=\mathbf{0}$, 因此根据 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 的线性无关性推出 $x_1=\cdots=x_n=0$. 所以 $\boldsymbol{\alpha}_1T,\cdots,\boldsymbol{\alpha}_nT$ 是线性无关的.

反过来, 假设向量 $\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1T,\cdots\boldsymbol{\beta}_n=\boldsymbol{\alpha}_nT$ 线性无关, 并回忆一下 6.2 节讲过的"变换 T 是一一映上当且仅当它有双边逆元素". 因为 V 是 n 维的, 所以 n 个线

性无关向量 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n$ 是 V 的一组基. 根据定理 1, 存在 V 的线性变换 S 使得

$$\boldsymbol{\beta}_1 S=\boldsymbol{\alpha}_1,\quad \boldsymbol{\beta}_2 S=\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\beta}_n S=\boldsymbol{\alpha}_n.$$

于是对每个 $i=1,\cdots,n$, 有 $\boldsymbol{\beta}_i(ST)=\boldsymbol{\beta}_i$. 因为 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n$ 是一组基, 所以根据定理 1 只存在一个线性变换 R, 使得对每个 i 有 $\boldsymbol{\beta}_i R=\boldsymbol{\beta}_i$, 于是 R 是恒等变换. 因此 $ST=I$. 类似地, 由于 $\boldsymbol{\alpha}_i(TS)=\boldsymbol{\beta}_i S=\boldsymbol{\alpha}_i$, 并且 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 是一组基, 所以 $TS=I$. 于是 S 是 T 的逆, 因而 T 是非奇异的.

这样, 为检验 T 是否是非奇异的, 我们可以检验 V 的任意有限基在 T 作用下的像的线性无关性, 检验的方法如 7.6 节所用的方法一样.

推论 1 设 T 是有限维向量空间 V 的线性变换. 如果 T 是非奇异的, 那么 (i) T 有双边线性逆; (ii) 由 $\boldsymbol{\xi}T=\mathbf{0}$ 和 $\boldsymbol{\xi}$ 在 V 中可推出 $\boldsymbol{\xi}=\mathbf{0}$; (iii) T 是从 V 到 V 的一一映射; (iv) T 把 V 变换到 V 上. 如果 T 是奇异的, 那么 (i′) T 既没有左逆也没有右逆; (ii′) 对某个 $\boldsymbol{\xi}\neq\mathbf{0}$ 有 $\boldsymbol{\xi}T=\mathbf{0}$; (iii′) T 不是一一的; (iv′) T 把 V 变换到 V 的某一真子空间中.

证明 条件 (i) 已在定理 9 中证明了. 再有, 如果对某个 $\boldsymbol{\xi}\neq\mathbf{0}$, 有 $\boldsymbol{\xi}T=\mathbf{0}$, 那么因为 $\mathbf{0}T=\mathbf{0}$, T 将不是一一的, 与非奇异的定义相矛盾. 这就说明 (ii). (iii) 和 (iv) 是定义的一部分. 其次, 如果 T 是奇异的, 那么对 V 的任意一组基 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$, 由定理 9, $\boldsymbol{\alpha}_1 T,\cdots,\boldsymbol{\alpha}_n T$ 线性相关. 因此对某一组不全为零的 $x_1,\cdots,x_n$ 有

$$\mathbf{0}=x_1\boldsymbol{\alpha}_1 T+\cdots+x_n\boldsymbol{\alpha}_n T=(x_1\boldsymbol{\alpha}_1+\cdots+x_n\boldsymbol{\alpha}_n)T=\boldsymbol{\xi}T.$$

因此 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 线性无关, 所以 $\boldsymbol{\xi}\neq\mathbf{0}$, 因此对某 $\boldsymbol{\xi}\neq\mathbf{0}$, 有 $\boldsymbol{\xi}T\neq\mathbf{0}$, 这就证明了 (ii′), 因为 $\mathbf{0}T=\mathbf{0}$, 所以由此推出 T 不是一一的, 因而证明了 (iii′), 再有, 因为 $\boldsymbol{\alpha}_1 T,\cdots,\boldsymbol{\alpha}_n T$ 线性相关, 而 V 是 n 维的, 所以它们张成 V 的某一真子空间, 根据 7.4 节定理 5 的推论 2, 这就证明了 (iv′), 最后, 根据 6.2 节定理 1, (iii′) 和 (iv′) 同 (i′) 是等价的. 证毕

注意, 因为推论 1 中列出的条件中每一对都是不相容的, 所以 8 个条件都是充分必要条件. 例如, 如果 (iv) 成立, 那么 (iv′) 就不能成立, 因此 T 不可能是奇异的, 所以它一定是非奇异的.

推论 2 如果有限维向量空间 V 的两个线性变换的乘积 TU 是恒等变换, 那么 T 和 U 两个都是非奇异的, 而且 $T=U^{-1}, U=T^{-1}, UT=I$.

证明 因为 $TU=I$, 所以 T 有右逆元素, 因此由上面的 (i′) 知 T 是非奇异的, 再根据 (i), T 有逆 T^{-1}, 那么有 $T^{-1}=T^{-1}(TU)=(T^{-1}T)U=U$, 如断言所述, 并且其他结论由此得出. 证毕

根据定理 6, F^n 的线性变换和 F 上的 $n\times n$ 矩阵之间在乘法之下是同构的, 所以上述结果可以平推到矩阵上去. 我们定义 $n\times n$ 矩阵 A 是非奇异的当且仅当在

定理 2 意义下它对应于 F^n 的一个非奇异线性变换 $T_{\boldsymbol{A}}$; 否则, 我们称 $\boldsymbol{A}$ 是奇异的. 而根据定理 2, 变换 $T_{\boldsymbol{A}}$ 把 F^n 的单位向量变到矩阵 $\boldsymbol{A}$ 的行, 因此定理 9 的条件变为 (参看 7.4 节定理 6 的推论)

推论 3 *域 F 上一个 $n \times n$ 矩阵是非奇异的当且仅当它的行是线性无关的, 或者等价于当且仅当这些行构成 F^n 的一组基.*

类似地, 推论 1 的条件 (i) 和 (i′) 可变为下面的结果.

推论 4 *一个 $n \times n$ 矩阵 $\boldsymbol{A}$ 是非奇异的当且仅当它有逆矩阵 $\boldsymbol{A}^{-1}$, 满足*

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I} \quad (\boldsymbol{A}, \boldsymbol{A}^{-1}, \boldsymbol{I} \text{ 都是 } n \times n \text{ 矩阵}). \tag{44}$$

如果 $\boldsymbol{A}$ 有逆, 则它的转置矩阵也有逆, 这是因为对 (44) 的两边取转置, 根据 (31) 我们得到 $(\boldsymbol{A}^{-1})^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}^{-1})^{\mathrm{T}} = \boldsymbol{I}$, 所以

$$(\boldsymbol{A}^{-1})^{\mathrm{T}} = (\boldsymbol{A}^{\mathrm{T}})^{-1}. \tag{45}$$

于是, 如果 $\boldsymbol{A}$ 是非奇异的, 则 $\boldsymbol{A}^{\mathrm{T}}$ 也是非奇异的; 而且反过来也有类似结论. 但是根据推论 3, $\boldsymbol{A}^{\mathrm{T}}$ 是非奇异的当且仅当它的行线性无关. 这些行实际上就是 $\boldsymbol{A}$ 的列, 因此我们有

推论 5 *一个方阵是非奇异的当且仅当它的列线性无关.*

如果把推论 2 从线性变换平推到矩阵上, 那么根据定理 6, 我们得到

推论 6 *每个方阵的左逆矩阵又是右逆矩阵.*

如果矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都有逆, 那么它们的乘积也有逆,

$$(\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1} \qquad (\text{注意反序!}) \tag{46}$$

这因为 $(\boldsymbol{A}\boldsymbol{B})(\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}) = \boldsymbol{A}(\boldsymbol{B}\boldsymbol{B}^{-1})\boldsymbol{A}^{-1} = \boldsymbol{A}\boldsymbol{I}\boldsymbol{A}^{-1} = \boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}$.

非奇异矩阵的逆可以通过求解适当的联立线性方程组来计算. 如果我们把基向量的坐标写成

$$\begin{aligned} \boldsymbol{I}_1 &= (1, 0, \cdots, 0), \\ \boldsymbol{I}_2 &= (0, 1, \cdots, 0), \\ &\vdots \\ \boldsymbol{I}_n &= (0, 0, \cdots, 1), \end{aligned} \tag{47}$$

那么在已知矩阵 $\boldsymbol{A} = (a_{ij})$ 中, 每一行 $\boldsymbol{A}_i$ 可写成这组基向量的线性组合

$$\boldsymbol{A}_i = \sum_j a_{ij}\boldsymbol{I}_j.$$

我们可以试图求解这组向量方程, 把 $\boldsymbol{I}_j$ 看成"未知向量", 把 $\boldsymbol{A}_i$ 看成已知向量; 解得 $\boldsymbol{I}_j$ 是线性表达式

$$\boldsymbol{I}_j = c_{j1}\boldsymbol{A}_1 + \cdots + c_{jn}\boldsymbol{A}_n = \sum_{k=1}^{n} c_{jk}\boldsymbol{A}_k. \tag{48}$$

根据 (40), 这个方程表明矩阵 $\boldsymbol{C} = (c_{jk})$ 满足 $\boldsymbol{CA} = \boldsymbol{I}$, 因此 $\boldsymbol{C} = \boldsymbol{A}^{-1}$. $\boldsymbol{A}^{-1}$ 的另一种构造方法将在 8.8 节中给出.

例 计算矩阵

$$\begin{pmatrix} 1 & 2 & -2 \\ -1 & 3 & 0 \\ 0 & -2 & 1 \end{pmatrix},$$

的逆, 把它的行写成 $\boldsymbol{A}_1 = \boldsymbol{I}_1 + 2\boldsymbol{I}_2 - 2\boldsymbol{I}_3, \boldsymbol{A}_2 = -\boldsymbol{I}_1 + 3\boldsymbol{I}_2, \boldsymbol{A}_3 = -2\boldsymbol{I}_2 + \boldsymbol{I}_3$. 这三个联立方程组有解

$$\begin{aligned} \boldsymbol{I}_1 &= 3\boldsymbol{A}_1 + 2\boldsymbol{A}_2 + 6\boldsymbol{A}_3 \\ \boldsymbol{I}_2 &= \boldsymbol{A}_1 + \boldsymbol{A}_2 + 2\boldsymbol{A}_3, \\ \boldsymbol{I}_3 &= 2\boldsymbol{A}_1 + 2\boldsymbol{A}_2 + 5\boldsymbol{A}_3. \end{aligned}$$

这些线性组合的所有系数 c_{jk} 组成一个逆矩阵, 因为我们可以验证

$$\begin{pmatrix} 3 & 2 & 6 \\ 1 & 1 & 2 \\ 2 & 2 & 5 \end{pmatrix}\begin{pmatrix} 1 & 2 & -2 \\ -1 & 3 & 0 \\ 0 & -2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

从无穷维向量空间 V 到第二个无穷维向量空间 W(在同一个域上) 的线性变换可以是一一的, 但不是映上, 反过来也是一样. 对于无穷维向量空间到它自身的线性变换, 上述结论同样正确. 例如, 无限实数序列的空间上的线性变换 $(x_1, x_2, x_3, \cdots) \longmapsto (0, x_1, x_2, x_3, \cdots)$ 是一一的, 但不是映上, 因此有很多 (线性) 右逆元素但没有左逆元素.

但是, 如果线性变换的双边逆存在, 那么即使 V 是无穷维空间, 这个逆也一定是线性的.

定理 10 *如果线性变换 $T: V \to W$ 是 V 到 W 上的一一变换, 那么它的逆也是线性的.*

证明 设 ψ 是 T 的唯一的逆变换, 是由 W 到 V 上的变换, 没有假定 ψ 是线性的. 在 W 中取向量 $\boldsymbol{\xi}$ 和 $\boldsymbol{\eta}$, 并取标量 c 和 d. 因为 ψT 是 W 的恒等变换, 而且 T 是线性的, 所以

$$(c\boldsymbol{\xi} + d\boldsymbol{\eta})\psi T = c\boldsymbol{\xi} + d\boldsymbol{\eta} = c(\boldsymbol{\xi}\psi T) + d(\boldsymbol{\eta}\psi T) = [c(\boldsymbol{\xi}\psi) + d(\boldsymbol{\eta}\psi)]T.$$

等式两边右乘 ψ, 因为 $T\psi$ 也是恒等变换, 所以我们到得

$$(c\boldsymbol{\xi}+d\boldsymbol{\eta})\psi=c(\boldsymbol{\xi}\psi)+d(\boldsymbol{\eta}\psi), \tag{49}$$

这个方程表明 ψ 是线性的. 证毕

在 7.8 节的意义之下, V 到 W 上的一一线性变换 T 是 V 到 W 的一个同构.

推论 1 *V 到 W 的同构 T 把 V 的任意一线性组无关向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 映射到 W 的一组线性无关向量, 并且把张成 V 的任意一组向量 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_s$ 映射到张成 W 的一组向量.*

证明 如果 $\boldsymbol{\alpha}_1T,\cdots,\boldsymbol{\alpha}_rT$ 之间存在线性关系

$$x_1(\boldsymbol{\alpha}_1T)+\cdots+x_r(\boldsymbol{\alpha}_rT)=\mathbf{0},$$

则我们可用 T^{-1} 右乘上式, 求得 $x_1\boldsymbol{\alpha}_1+\cdots+x_r\boldsymbol{\alpha}_r=\mathbf{0}$, 因此 $x_1=\cdots=x_r=0$, 从而 $\boldsymbol{\alpha}_1T,\cdots,\boldsymbol{\alpha}_rT$ 线性无关. 后一半的证明类似.

对于任意变换 $T:V\to W$, V 的子空间 S 在 T 之下的像即变换式 $S'=(S)T$ 被定义为 S 中所有向量 $\boldsymbol{\xi}$ 的变换式 $\boldsymbol{\xi}T$ 的集合. 这个像总是 W 的子空间, 因为 S' 中向量 $\boldsymbol{\xi}T$ 和 $\boldsymbol{\eta}T$ 的每个线性组合 $c(\boldsymbol{\xi}T)+d(\boldsymbol{\eta}T)=(c\boldsymbol{\xi}+d\boldsymbol{\xi})T$ 仍在 S' 中.

推论 2 *对于同构 $T:V\to W$, V 的任意有限维子空间 S 在 T 之下的像的维数同 S 的维数一样. 于是, T 把直线映射到直线, 把平面映射到平面.*

习 题

1. 求出 8.3 节习题 1 中矩阵 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{D}$ 的逆.
2. (a) 证明: $\boldsymbol{A}=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 是非奇异的当且仅当 $ad-bc\neq 0$.

   (b) 证明: 如果 $\boldsymbol{A}$ 是非奇异的, 那么它的逆是 $\Delta^{-1}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$, 这里 $\Delta=ad-bc$.
3. 求出 8.1 节中线性变换 R_θ, D_k, S_a 的逆.
4. 求出 8.1 节习题 4 中的线性变换的逆.
5. (a) 当 $\theta=45°$, 计算变换 $R_\theta^{-1}U_bR_\theta$(见 8.1 节) 对应的矩阵, 这里 U_b 是变换 $x'=bx, y'=y$.

   (b) 描述这个变换的几何意义.

   (c) 对变换 $R_\theta^{-1}S_aR_\theta$(其中 $\theta=45°$), 做 (a) 和 (b).
6. 证明: 如果 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2-\boldsymbol{A}+\boldsymbol{I}=\boldsymbol{O}$, 那么 $\boldsymbol{A}^{-1}$ 存在, 并且等于 $\boldsymbol{I}-\boldsymbol{A}$.
7. 求出 8.3 节习题 9 中矩阵 $\boldsymbol{E}_1,\boldsymbol{E}_2$ 和 E_3 的逆.
8. 求出 8.5 节习题 3 中矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的逆. (提示: 用分块矩阵.)
9. (a) 给出 2×2 三角形矩阵的逆矩阵计算公式.

(b) 给出 3×3 三角形矩阵的逆矩阵计算公式. (提示: 用三角形逆矩阵试一试.)

(c) 证明: 对角线上的元素不为零的每个三角形矩阵的有一个三角形的逆矩阵.

10. 已知 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{A}^{-1},\boldsymbol{B}^{-1},\boldsymbol{C}$, 求出下列矩阵的逆,

(a) $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B} \end{pmatrix}$, (b) $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{O} & \boldsymbol{B} \end{pmatrix}$, (c) $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{C} & \boldsymbol{B} \end{pmatrix}$.

11. 证明: 所有非奇异的 $n\times n$ 矩阵关于矩阵乘法构成群.

12. 证明: 如果方阵的乘积 $\boldsymbol{AB}$ 是非奇异的, 那么它的两个因子 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 也是非奇异的.

*13. 不用线性变换来证明, 矩阵 $\boldsymbol{A}$ 有左逆矩阵当且仅当 $\boldsymbol{A}$ 的行线性无关.

14. 证明定理 10 的推论 2.

15. 列出一个序列 $(x_1,x_2,x_3,\cdots)$ 的空间到它自身上的线性变换, 它不是一一的.

*16. 证明: 如果线性变换 $T:V\to W$ 有右逆元素, 那么它的右逆元素也是线性的 (没有假定是有限维的).

8.7 秩与零度

在一般情况下 (见 6.2 节), 每个变换 (函数)$T:S\to S_1$, 以 S 为定义域, 以 S_1 为取值域. T 的值域是变换式的集合 (即定义域在 T 之下的像).

当 T 是向量空间 V 到另一个向量空间 W 的线性变换时, V 的像 (所有 $\boldsymbol{\xi}T$ 的集合) 不可能是 W 的任意子集合.

引理 1 线性变换$T:V\to W$ 的像本身是一个向量空间 (因此是 W 的子空间).

证明 因为 $c(\boldsymbol{\xi}T)=(c\boldsymbol{\xi})T,\boldsymbol{\xi}T+\boldsymbol{\eta}T=(\boldsymbol{\xi}+\boldsymbol{\eta})T$, 所以变换式的集合在向量加法和数乘运算之下是封闭的.

引理 2 设 $T_{\boldsymbol{A}}$ 是对应于 $m\times n$ 矩阵 $\boldsymbol{A}$ 的线性变换, 那么 $T_{\boldsymbol{A}}$ 的像是 $\boldsymbol{A}$ 的行空间.

证明 变换 $T_{\boldsymbol{A}}:F^m\to F^n$ 把 F^m 的每个向量 $\boldsymbol{X}=(x_1,\cdots,x_m)$ 映射到 F^n 中的 $\boldsymbol{Y}=\boldsymbol{XA}$, 所以 $T_{\boldsymbol{A}}$ 的像是由所有形为

$$\boldsymbol{Y}=\boldsymbol{XA}=\left(\sum x_i a_{i1},\cdots,\sum x_i a_{in}\right)=\sum x_i(a_{i1},\cdots,a_{in})$$

的 n- 数组构成. 这恰好就是 $\boldsymbol{A}$ 的行 $\boldsymbol{A}_i=(a_{i1},\cdots,a_{in})$ 的全体不同线性组合. 于是 $T_{\boldsymbol{A}}$ 的值域是由 $\boldsymbol{A}$ 的行向量的一切线性组合组成的集合. 正如 7.5 节中所定义的, 这就是 $\boldsymbol{A}$ 的行空间. 证毕

7.6 节中我们已经定义矩阵 $\boldsymbol{A}$ 的秩为 $\boldsymbol{A}$ 的行空间的 (线性) 维数, 因此它也是 $T_{\boldsymbol{A}}$ 的值域的维数. 更一般地, 任意线性变换 T 的秩定义为 T 的像的维数 (有限或无限).

因为由 m 个已知向量张成的子空间的维数等于这个子空间中线性无关向量的最大个数, 所以 $\boldsymbol{A}$ 的秩也等于 $\boldsymbol{A}$ 的线性无关行向量的最大个数. 由于这个理由, $\boldsymbol{A}$

按上面定义的秩常常称为 A 的行秩，它不同于列秩，列秩是 $\boldsymbol{A}$ 的线性无关列向量的最大个数.

同矩阵的行空间或者线性变换的值域概念成对偶的是它的零空间概念.

定义 线性变换 T 的零空间是使得 $\boldsymbol{\xi}T=\mathbf{0}$ 的所有向量 $\boldsymbol{\xi}$ 的集合. 矩阵 $\boldsymbol{A}$ 的零空间是满足齐次线性方程组 $\boldsymbol{XA}=\mathbf{0}$ 的所有行矩阵 $\boldsymbol{X}$ 的集合.

引理 3 任意线性变换 (或矩阵) 的零空间是它的定义域的一个子空间.

证明 如果 $\boldsymbol{\xi}T=\mathbf{0}$ 和 $\boldsymbol{\eta}T=\mathbf{0}$, 那么对所有 c 和 d, 有

$$(c\boldsymbol{\xi}+d\boldsymbol{\eta})T=c(\boldsymbol{\xi}T)+d(\boldsymbol{\eta}T)=\mathbf{0}+\mathbf{0}=\mathbf{0},$$

因此 $c\boldsymbol{\xi}+d\boldsymbol{\eta}$ 在零空间中，所以零空间是一个子空间. 证毕

已知矩阵 $\boldsymbol{A}$ 或线性变换 T 的零空间的维数称为 $\boldsymbol{A}$ 或 T 的零度(nullity). 零度和秩的关系满足一个对于矩阵和线性变换都成立的基本方程. 因为矩阵和线性变换之间存在对应关系，我们只须对一种情况给出证明.

定理 11 秩与零度之和等于定义域的维数.

比如，对于 $m\times n$ 矩阵，(行) 秩与 (行) 零度之和等于 m.

证明 如果 T 的零度是 s, 那么 T 的零空间 N 有 s 个元素构成的基 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s$, 可以把它扩充成 T 的整个定义域的基 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$. 因为每个 $\boldsymbol{\alpha}_iT=\mathbf{0}$, 所以 $\boldsymbol{\beta}_1T,\cdots,\boldsymbol{\beta}_rT$ 张成 T 的像 R. 此外，由 $x_1(\boldsymbol{\beta}_1T)+\cdots+x_r(\boldsymbol{\beta}_rT)=\mathbf{0}$ 推出 $x_1\boldsymbol{\beta}_1+\cdots+x_r\boldsymbol{\beta}_r$ 在 N 中，所以 $x_1=\cdots=x_r=0$. 因此向量 $\boldsymbol{\beta}_1T,\cdots,\boldsymbol{\beta}_rT$ 线性无关，并构成 R 的基. 我们得出结论：T 的定义域的维数 m 是 N 的维数 s 与 R 的维数 r 之和：$m=s+r$. 这就是我们要证的.

定理 12 一个线性变换 $T:F^n\to F^n$ 是非奇异的充分必要条件是下面条件之一：

(a) T 的秩等于 n; (b) T 的零度等于 0.

证明 条件 (a) 表明，T 把 F^n 映上到它自身；而条件 (b) 表明，$\boldsymbol{\xi}T=\mathbf{0}$ 可推出 $\boldsymbol{\xi}=\mathbf{0}$ 在 F^n 中. 这样，定理 12 正好重新叙述了定理 9 的推论 1 中的条件 (iv) 和条件 (ii).

习 题

1. 求出 8.1 节习题 1(a)~(d), 习题 4(a), (b) 中所给出的线性变换的值域、零空间、秩以及零度.
2. 构造一个由 $\mathbf{R}^3$ 到它自身的线性变换，使得它的值域由向量 (1, 3, 2) 和 (3, −1, 1) 张成.
3. 构成一个由 $\mathbf{R}^4$ 到它自身的线性变换，使得它的零空间由向量 (1, 2, 3, 4) 和 (2, 2, 4, 4) 张成.
4. 证明：乘积 $\boldsymbol{AB}$ 的行秩决不能超过 $\boldsymbol{A}$ 的行秩.

5. 证明：如果 $n \times n$ 矩阵 $\boldsymbol{A}$ 是非奇异的，那么对每个 $n \times n$ 矩阵 $\boldsymbol{B}$，矩阵 $\boldsymbol{AB}, \boldsymbol{B}$ 和 $\boldsymbol{BA}$ 都具有相同的秩.
6. 证明：$\operatorname{rank}(\boldsymbol{A}+\boldsymbol{B}) \leqslant \operatorname{rank}(\boldsymbol{A})+\operatorname{rank}(\boldsymbol{B})$.
7. 如果已知 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的秩，那么矩阵 $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B} \end{pmatrix}$ 的秩等于多少？

8.8 初等矩阵

在 7.5 节中我们引进作用在矩阵 $\boldsymbol{A}$ 上的初等行运算，它们可以解释为用适当的矩阵左乘矩阵 $\boldsymbol{A}$. 例如，矩阵中的两行互换，可以用一个矩阵左乘这个矩阵来实现，乘上的这个矩阵是把单位矩阵 $\boldsymbol{I}$ 相应的行互换而得到的. 例如，

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}=\begin{pmatrix} 0\cdot a_1+1\cdot b_1 & 0\cdot a_2+1\cdot b_2 \\ 1\cdot a_1+0\cdot b_1 & 1\cdot a_2+0\cdot b_2 \end{pmatrix}=\begin{pmatrix} b_1 & b_2 \\ a_1 & a_2 \end{pmatrix}.$$

为了把矩阵的第二行加到第一行上去或者把第二行乘以标量 c，只须对矩阵前面的单位矩阵因子做同样的运算：

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}=\begin{pmatrix} a_1+b_1 & a_2+b_2 \\ b_1 & b_2 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix}\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}=\begin{pmatrix} a_1 & a_2 \\ cb_1 & cb_2 \end{pmatrix}.$$

对 $m \times n$ 矩阵的情形，类似的结果也成立. 用来表示这些初等行运算的左乘因子称为初等矩阵.

定义 对 $m \times m$ 单位矩阵 $\boldsymbol{I}$ 做一次初等行运算所得到的矩阵 $\boldsymbol{E}$ 称为 $m \times m$ 初等矩阵.

于是有三类初等矩阵，它们的样本是

$$\underset{\boldsymbol{H}_{24}}{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}},\quad \underset{\boldsymbol{I}+2\boldsymbol{E}_{33}}{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}},\quad \underset{\boldsymbol{I}+d\boldsymbol{E}_{21}}{\begin{pmatrix} 1 & 0 & 0 & 0 \\ d & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}. \tag{50}$$

一般地，设 $\boldsymbol{I}_k$ 表示 $m \times m$ 单位矩阵 $\boldsymbol{I}$ 的第 k 行，那么把 $\boldsymbol{I}$ 中的第 i 行与第 j 行互换，得到初等置换矩阵 $\boldsymbol{H}=(h_{ij})$，它的各行 $\boldsymbol{H}_k$ 是

$$\boldsymbol{H}_i=\boldsymbol{I}_j,\quad \boldsymbol{H}_j=\boldsymbol{I}_i,\quad \boldsymbol{H}_k=\boldsymbol{I}_k \quad (k \neq i,j). \tag{51}$$

类似地, $\boldsymbol{I}$ 的第 i 行乘上一个非零标量 c, 得到矩阵 $\boldsymbol{M}$, 它的各行 $\boldsymbol{M}_k$ 是

$$\boldsymbol{M}_i = c\boldsymbol{I}_i(c \neq 0), \qquad \boldsymbol{M}_k = \boldsymbol{I}_k(k \neq i). \tag{52}$$

如果 $\boldsymbol{E}_{ij}$ 如前所述, 是一个只在第 i 行第 j 列上有一个元素 1, 其他所有元素都为 0 的矩阵, 那么矩阵 $\boldsymbol{M}$ 可以写成 $\boldsymbol{M} = \boldsymbol{I} + (c-1)\boldsymbol{E}_{ij}$. 最后, 把 $\boldsymbol{I}$ 的第 i 行乘上 d 后加到第 j 行上去, 得到初等矩阵 $\boldsymbol{F} = \boldsymbol{I} + d\boldsymbol{E}_{ji}$, 它的各行 $\boldsymbol{F}_k$ 是

$$\boldsymbol{F}_j = \boldsymbol{I}_j + d\boldsymbol{I}_i, \quad \boldsymbol{F}_k = \boldsymbol{I}_k(k \neq j). \tag{53}$$

定理 13 *对 $m \times n$ 矩阵 $\boldsymbol{A}$ 所做的每种初等行运算, 相当于对矩阵 $\boldsymbol{A}$ 左乘一个相应的 $m \times m$ 初等矩阵 $\boldsymbol{E}$.*

通过直接计算乘积 $\boldsymbol{EA}$, 我们可以容易地得到定理的证明. 例如, 考虑把 $\boldsymbol{A}$ 的第 i 行加到 $\boldsymbol{A}$ 的第 j 行上去的初等行运算. 相应的初等矩阵 $\boldsymbol{F}$ 的各行 $\boldsymbol{F}_k$ 由 (53) 给出, 乘积 $\boldsymbol{FA}$ 的每一行总是由第一个因子的各行按公式 (37) 求出, 所以

$$\begin{aligned}(\boldsymbol{FA})_j &= \boldsymbol{F}_j\boldsymbol{A} = (\boldsymbol{I}_i + \boldsymbol{I}_j)\boldsymbol{A} = \boldsymbol{I}_i\boldsymbol{A} + \boldsymbol{I}_j\boldsymbol{A} = (\boldsymbol{IA})_i + (\boldsymbol{IA})_j,\\ (\boldsymbol{FA})_k &= \boldsymbol{F}_k\boldsymbol{A} = \boldsymbol{I}_k\boldsymbol{A} = (\boldsymbol{IA})_k \qquad (k \neq j).\end{aligned}$$

这些方程表明, $\boldsymbol{FA}$ 的各行是把 $\boldsymbol{IA} = \boldsymbol{A}$ 的第 i 行加到第 j 行上得到的. 换句话说, 这里所讨论的初等行运算把 $\boldsymbol{A}$ 变为 $\boldsymbol{FA}$, 正如定理 13 所断言的.

推论 1 *每个初等矩阵 $\boldsymbol{E}$ 是非奇异的.*

证明 $\boldsymbol{E}$ 是从 $\boldsymbol{I}$ 通过某些行运算得到的. 这些运算的反运算对应着某个初等矩阵 $\boldsymbol{E}^*$, 并把 $\boldsymbol{E}$ 变回到 $\boldsymbol{I}$. 根据定理 13, 它把 $\boldsymbol{E}$ 变为 $\boldsymbol{E}^*\boldsymbol{E}$ 所以 $\boldsymbol{E}^*\boldsymbol{E} = \boldsymbol{I}$, 因而 $\boldsymbol{E}$ 有左逆矩阵 $\boldsymbol{E}^*$, 所以是非奇异的.

推论 2 *如果两个 $m \times n$ 矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是行等价的, 那么 $\boldsymbol{B} = \boldsymbol{PA}$, 其中 $\boldsymbol{P}$ 是非奇异矩阵.*

这因为, 根据定理 13 有, $\boldsymbol{B} = \boldsymbol{E}_n\boldsymbol{E}_{n-1}\cdots\boldsymbol{E}_1\boldsymbol{A}$, 其中每个 $\boldsymbol{E}_i$ 都是初等矩阵, 所以 $\boldsymbol{P}$ 是非奇异的.

初等行运算和用初等矩阵左乘这两种运算之间的等价性对高斯消去法给出另一个有用的解释. 在通常情况下, 最后在主对角线上没有出现零, 在这种情况下, 一方面系数矩阵 $\boldsymbol{A}$ 被化成上三角形矩阵 U(这是显然的), 另一方面, 因为从后面的行减去第 i 行的倍数的运算相当于用一个下三角形矩阵 $\boldsymbol{L}_k$ 左乘, 所以我们有

$$\boldsymbol{U} = \boldsymbol{L}_s\boldsymbol{L}_{s-1}\cdots\boldsymbol{L}_1\boldsymbol{A} = \boldsymbol{LA}, \qquad s \leqslant \frac{n(n-1)}{2},$$

这里 $\boldsymbol{L} = \boldsymbol{L}_s\boldsymbol{L}_{s-1}\cdots\boldsymbol{L}_1$ 是下三角形矩阵. 因此 $\boldsymbol{AX} = \boldsymbol{B}$ 等价于 $\boldsymbol{UX} = \boldsymbol{LB}$, 这里 $\boldsymbol{U} = \boldsymbol{LA}$. 因此我们可以写成 $\boldsymbol{A} = \boldsymbol{L}^{-1}\boldsymbol{U}$, 这里 $\boldsymbol{L}^{-1}$ 也是下三角形矩阵而 $\boldsymbol{U}$ 是上三角形矩阵, 这称为 $\boldsymbol{A}$ 的“$\boldsymbol{LU}$ 分解”.

矩阵的逆可以用初等矩阵来计算. 设 $\boldsymbol{A}$ 是任意非奇异方阵, 根据 7.6 节定理 9 的推论 3, 可以通过初等行运算把 $\boldsymbol{A}$ 化为单位矩阵 $\boldsymbol{I}$. 因此根据定理 13, 对适当的初等矩阵 $\boldsymbol{E}_1,\cdots,\boldsymbol{E}_s$, 我们有

$$\boldsymbol{E}_s\boldsymbol{E}_{s-1}\cdots\boldsymbol{E}_1\boldsymbol{A}=\boldsymbol{I}.$$

这个方程的两边都右乘 $\boldsymbol{A}^{-1}$, 那么有

$$\boldsymbol{E}_s\boldsymbol{E}_{s-1}\cdots\boldsymbol{E}_1\boldsymbol{I}=\boldsymbol{A}^{-1}. \tag{54}$$

等式左边的矩阵是这列初等运算 $\boldsymbol{E}_1,\cdots,\boldsymbol{E}_s$ 作用到单位矩阵 $\boldsymbol{I}$ 上而得到的结果. 这就证明了

定理 14 如果一个方阵 $\boldsymbol{A}$ 通过一系列行运算化为单位矩阵 $\boldsymbol{I}$, 那么把这同一系列运算作用到单位矩阵 $\boldsymbol{I}$ 上, 就给出矩阵 $\boldsymbol{A}$ 的逆矩阵.

这是求逆矩阵的一个有效方法. 给定任意矩阵 $\boldsymbol{A}$, 通过有限次有理运算或者得出 $\boldsymbol{A}$ 的逆矩阵, 或者化成一个等价的奇异矩阵. 后一种情况 $\boldsymbol{A}$ 没有逆. 对于大于 3×3 的矩阵 $\boldsymbol{A}$, 这种方法比起用行列式理论求逆 $\boldsymbol{A}^{-1}$(参看第 10 章) 更为有效.

附带说一下, 任意非奇异矩阵 $\boldsymbol{P}$ 是另一非奇异矩阵的逆 $(\boldsymbol{P}^{-1})^{-1}$, 因此, 像 (54) 中表示的那样, $\boldsymbol{P}$ 可写成初等矩阵的乘积. 这同定理 13 的推论 1 结合起来得出下面结果.

定理 15 方阵 $\boldsymbol{P}$ 是非奇异的当且仅当它可以表示成初等矩阵的乘积

$$\boldsymbol{P}=\boldsymbol{E}_s\boldsymbol{E}_{s-1}\cdots\boldsymbol{E}_1. \tag{55}$$

推论 1 两个 $m\times n$ 矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是行等价的当且仅当 $\boldsymbol{B}=\boldsymbol{P}A$, 其中 $\boldsymbol{P}$ 是某一非奇异矩阵.

因为 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 行等价当且仅当 $\boldsymbol{B}=\boldsymbol{E}_n\boldsymbol{E}_{n-1}\cdots\boldsymbol{E}_1\boldsymbol{A}$ 其中 $\boldsymbol{E}_1,\cdots,\boldsymbol{E}_n$ 都是初等矩阵 (定理 13). 再根据定理 15, 这相当于 $\boldsymbol{B}=\boldsymbol{P}A$, 其中 $\boldsymbol{P}$ 是非奇异的.

定理 15 在二维情形下有简单的几何解释. 2×2 初等矩阵只有下面几种

$$\boldsymbol{H}_{12}=\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad \boldsymbol{M}_1=\begin{pmatrix}c&0\\0&1\end{pmatrix},\quad \boldsymbol{M}_2=\begin{pmatrix}1&0\\0&c\end{pmatrix},$$

$$\boldsymbol{F}_{12}=\begin{pmatrix}1&0\\d&1\end{pmatrix},\quad \boldsymbol{F}_{21}=\begin{pmatrix}1&d\\0&1\end{pmatrix}.$$

相对应的线性变换, 如 8.1 节中给出的那样, 是

$\boldsymbol{H}_{12}$	是对于过原点、与 x 轴成 45° 角的直线的平面反射.
$\boldsymbol{M}_i(c>0)$	是平行于 x 轴或 y 轴的压缩 (或伸长).
$\boldsymbol{M}_i(c<0)$	是先压缩 (或伸长), 然后再对于 x 轴或 y 轴进行反射.
$\boldsymbol{F}_{ij}$	是平行于某一轴的切变换.

这就得到

推论 2 *平面上任意非奇异齐次线性变换可以表示成切变换、一维压缩 (或伸长) 及反射的乘积.*

这一基本的几何结论是通过矩阵进行代数论证而得到的. 对于三维或更高维空间可以得出类似的结果.

矩阵的初等行运算只包含已知域 $\boldsymbol{F}$ 内的运算. 如果矩阵 $\boldsymbol{A}$ 的元素都是有理数, 而所考虑的域是实数域, 那么初等运算可以同只包含有理数的域一样进行. 在这两个域中, 我们得到相同的梯形矩阵, 因此线性无关的行的个数相同.

定理 16 *如果域 $\boldsymbol{F}$ 上的矩阵 A 的所有元素都属于一个比 $\boldsymbol{F}$ 小的域 $\boldsymbol{F}'$, 那么 $\boldsymbol{A}$ 相对于域 $\boldsymbol{F}$ 的秩同 $\boldsymbol{A}$ 相对于域 $\boldsymbol{F}'$ 的秩一样.*

行的等价运算恰好可用来解联立线性方程组(2.3 节和 7.5 节). 为了描述它们之间的联系, 我们考虑 n 个未知数 $x_1, \cdots, x_n$ 的 m 个方程

$$\sum_j a_{ij}x_j = b_i \quad (i = 1, \cdots, m;\ j = 1, \cdots, n).$$

未知数的全体系数构成 $m \times n$ 矩阵 $\boldsymbol{A} = (a_{ij})$, 而常数项 $b_1, \cdots, b_m$ 构成列向量 $\boldsymbol{B}^{\mathrm{T}}$. 方程组可以写成矩阵形式 $\boldsymbol{AX}^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}$, 其中 $\boldsymbol{X}^{\mathrm{T}}$ 是未知数的列向量 (是行向量 $\boldsymbol{X} = (x_1, \cdots, x_n)$ 的转置). 常数列向量 $\boldsymbol{B}^{\mathrm{T}}$ 可以添加到已知系数矩阵 $\boldsymbol{A}$ 中构成 $m \times (n+1)$ 矩阵 $(\boldsymbol{A}\ \boldsymbol{B}^{\mathrm{T}})$, 这就是所谓的已知方程组的*增广矩阵*. 对增广矩阵的行进行的运算, 对应于把已知方程组化为等价方程组的运算, 所以,*如果两个方程组 $\boldsymbol{AX}^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}$和$\boldsymbol{A}^*\boldsymbol{X}^{\mathrm{T}} = \boldsymbol{B}^{*\mathrm{T}}$ 的增广矩阵是行等价的, 那么这两个方程组有相同的解 $\boldsymbol{X}^{\mathrm{T}}$.*

习　题

1. 对 8.3 节习题 9(a) 中列出的矩阵, 求出它们的行等价梯形矩阵.
2. (a) 列出所有可能的 3×3 初等矩阵.
   (b) 画图表示每个形如 (51) ~ (53) 的 $n \times n$ 初等矩阵.
3. 分别求出正文中列出的 4×4 初等矩阵 $\boldsymbol{H}_{24}, \boldsymbol{I} + 2\boldsymbol{E}_{33}, \boldsymbol{I} + d\boldsymbol{E}_{21}$ 的逆.
4. 通过对定理 15 下面的 5 个矩阵直接进行计算, 证明 2×2 矩阵情形下的定理 13.
5. 求出矩阵

$$\begin{pmatrix} 1 & 0 & 3 \\ 2 & 4 & 1 \\ 1 & 3 & 0 \end{pmatrix} \text{和} \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix} \text{的逆.}$$

6. 把下列各矩阵表示成初等矩阵的乘积:

(a) $\begin{pmatrix} 3 & 6 \\ 2 & 1 \end{pmatrix}$, (b) $\begin{pmatrix} 4 & -2 \\ 3 & -5 \end{pmatrix}$, (c) 习题 5 的第一个矩阵.

7. 把变换 $x' = 2x - 5y, y' = -3x + y$ 表示成切变换、一维压缩及反射的乘积.

*8. 对三维空间叙述并证明与定理 15 的推论 2 相类似的命题. 用 7.5 节习题 3 改进你的结果.

9. 证明：任意 2×2 非奇异矩阵可以表示成矩阵

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{和} \begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix} (c \neq 0 \text{ 为任意标量})$$

的乘积. 这个结果的几何意义是什么?

10. 证明：矩阵乘积的秩绝不能超过它的因子中任何一个的秩.

11. 证明：线性方程组 $\boldsymbol{A}\boldsymbol{X}^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}$ 有解当且仅当矩阵 $\boldsymbol{A}$ 的秩等于增广矩阵 $(\boldsymbol{A}\ \boldsymbol{B}^{\mathrm{T}})$ 的秩.

12. 设非齐次线性方程组 $\boldsymbol{A}\boldsymbol{X}^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}$ 的特解为 $\boldsymbol{X}^{\mathrm{T}} = \boldsymbol{X}_0^{\mathrm{T}}$, 证明：该方程组的每一个解 $\boldsymbol{X}^{\mathrm{T}}$ 可以表示成 $\boldsymbol{X}^{\mathrm{T}} = \boldsymbol{X}_0^{\mathrm{T}} + \boldsymbol{Y}^{\mathrm{T}}$, 其中 $\boldsymbol{Y}^{\mathrm{T}}$ 是齐次线性方程组 $\boldsymbol{A}\boldsymbol{Y}^{\mathrm{T}} = 0$ 的解. 反之亦然.

13. 证明：如果系数在域 F 中的线性方程组在 F 中没有解, 那么它在任意比 F 大的域中也没有解.

8.9 等价与标准型

类似于初等行运算的运算也可以作用到矩阵的列上. 于是作用到 $m \times n$ 矩阵 $\boldsymbol{A}$ 上的*初等列运算*是指下面运算中的任何一种：(i) $\boldsymbol{A}$ 的任意两列互换; (ii) 任意一列乘以一个非零标量; (iii) 某一列的任意倍数加到另一列上.

如果用转置矩阵 $\boldsymbol{A}^{\mathrm{T}}$ 代替矩阵 $\boldsymbol{A}$, 可以把初等列运算变成初等行运算, 反之亦然. 特别是, $\boldsymbol{A}$ 可以通过一系列初等列运算变换成 $\boldsymbol{B}$ 当且仅当 $\boldsymbol{A}^{\mathrm{T}}$ 可以通过一系列初等行运算变换成 $\boldsymbol{B}^{\mathrm{T}}$. 根据定理 15 的推论 1, 这就意味着 $\boldsymbol{B}^{\mathrm{T}} = \boldsymbol{P}\boldsymbol{A}^{\mathrm{T}}$ 或者 $\boldsymbol{B} = (\boldsymbol{B}^{\mathrm{T}})^{\mathrm{T}} = (\boldsymbol{P}\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{A}\boldsymbol{P}^{\mathrm{T}} = \boldsymbol{A}\boldsymbol{Q}$, 其中 $\boldsymbol{Q} = \boldsymbol{P}^{\mathrm{T}}$ 是非奇异矩阵. 反过来, $\boldsymbol{B} = \boldsymbol{A}\boldsymbol{Q}$, 这里非奇异矩阵 $\boldsymbol{Q}$ 使得 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 列等价. 因此, 列运算作用到矩阵上等价于用非奇异矩阵右乘这个矩阵. 同每个初等列运算相对应的右乘因子, 可以通过把这个初等列运算作用到单位矩阵上来求出, 差不多和定理 13 一样.

初等列运算与初等行运算可以一起使用. 我们可以定义两个 $m \times n$ 矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$等价当且仅当 $\boldsymbol{A}$ 通过一系列初等行运算和初等列运算变换成 $\boldsymbol{B}$, 于是我们得到下面结果.

定理 17 *两个 $m \times n$ 矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 等价当且仅当对适当的 $m \times m$ 非奇异矩阵 $\boldsymbol{P}$ 和 $n \times n$ 非奇异矩阵 $\boldsymbol{Q}$, 有 $\boldsymbol{B} = \boldsymbol{P}\boldsymbol{A}\boldsymbol{Q}$.*

联合使用初等行运算和初等列运算, 我们可以把矩阵化成非常简单的标准型(见 9.5 节).

定理 18 任意 $m\times n$ 矩阵 $\boldsymbol{A}$ 等价于一个对角矩阵 $\boldsymbol{D}$, 其中对角线元素 (是指元素的行标和列标相同) 或者是 0 或者是 1, 并且在对角线上所有的 1 在所有的 0 前面.

显然, 如果 r 是 $\boldsymbol{D}$ 中非零元素的个数, 当然, $r\leqslant m, r\leqslant n$, 那么 $\boldsymbol{D}=\boldsymbol{D}_r$ 可以表示成分块矩阵为

$$\boldsymbol{D}_r=\begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{O}_{r,n-r} \\ \boldsymbol{O}_{m-r,r} & \boldsymbol{O}_{m-r,n-r} \end{pmatrix}, \tag{56}$$

其中 $\boldsymbol{I}_r$ 是 $r\times r$ 单位矩阵, $\boldsymbol{O}_{i,j}$ 表示 $i\times j$ 零矩阵.

通过对 $\boldsymbol{A}$ 的行数 m 使用归纳法, 来证明这个定理. 如果 $\boldsymbol{A}$ 的所有元素都是零, 那就无须证明. 若不然, 通过行置换与列置换, 我们可以把某个非零元素 c 移到 a_{11} 的位置. 然后第一行乘上 c^{-1}, a_{11} 位置的元素就化为 1. 再分别把第一行乘上适当的倍数加到其他各行中, 可以把第一列的其他元素都化为零. 用同样的方法可以把第一行的其他元素化为零. 于是矩阵 $\boldsymbol{A}$ 就化为下面形式的等价矩阵:

$$\boldsymbol{B}=\begin{pmatrix} 1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{C} \end{pmatrix},\ \boldsymbol{C}\text{ 是}(m-1)\times(n-1)\text{矩阵}. \tag{57}$$

再根据对矩阵 $\boldsymbol{C}$ 做的归纳法假定, 就证明了这个定理.

定理 19 等价矩阵具有相同的秩.

证明 我们已经知道 (7.5 节定理 7), 行等价矩阵具有相同的行空间, 因而具有相同的秩. 因此我们只须证明列等价矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}=\boldsymbol{AQ}$($\boldsymbol{Q}$ 是非奇异矩阵) 具有相同秩. 再有, 根据定理 11, 如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 具有相同的零度, 则上述结论正确, 当 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 有相同的零空间时, 它们一定具有相同的零度. 事实上, 由 $\boldsymbol{XA}=\boldsymbol{O}$ 显然可推出 $\boldsymbol{XB}=\boldsymbol{XAQ}=\boldsymbol{OQ}=\boldsymbol{O}$; 反过来, 由 $\boldsymbol{XB}=\boldsymbol{O}$ 可推出 $\boldsymbol{XA}=\boldsymbol{XAQQ}^{-1}=\boldsymbol{XBQ}^{-1}=\boldsymbol{OQ}^{-1}=\boldsymbol{O}$, 这就是说, 列等价矩阵具有相同的零空间.

推论 1 一个 $m\times n$ 矩阵 $\boldsymbol{A}$ 同一个且只同一个形为 (56) 的对角矩阵等价; $\boldsymbol{A}$ 的秩 r 由对角线上的 1 的个数 r 来确定.

推论 2 等价矩阵具有相同的列秩.

证明 矩阵 $\boldsymbol{A}$ 的列秩 ($\boldsymbol{A}$ 的线性无关列向量的最大个数) 等于 $\boldsymbol{A}$ 的转置矩阵 $\boldsymbol{A}^{\mathrm{T}}$ 的行秩. 但是 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的等价性推出 $\boldsymbol{A}^{\mathrm{T}}$ 和 $\boldsymbol{B}^{\mathrm{T}}$ 的等价性. 根据定理 19, $\boldsymbol{A}^{\mathrm{T}}$ 和 $\boldsymbol{B}^{\mathrm{T}}$ 具有相同的秩, 所以 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 具有相同的列秩.

标准型 (56) 中矩阵的秩同列秩一样; 根据等价性, 秩是不变的, 所以我们导出

推论 3 矩阵的 (行) 秩总等于它的列秩.

推论 4 两上 $m\times n$ 矩阵是等价的当且仅当它们具有相同的秩.

如果两个矩阵等价, 则它们具有相同的秩 (定理 19); 如果两个矩阵有相同的秩, 则两个矩阵都等价于同一个标准型 $\boldsymbol{D}$(推论 1), 因此它们彼此等价.

推论 5 $n \times n$ 矩阵 $\boldsymbol{A}$ 是非奇异的当且仅当它与单位矩阵 $\boldsymbol{I}$ 等价.

这是因为, 由推论 4, $\boldsymbol{A}$ 等价于 $\boldsymbol{I}$ 当且仅当 $\boldsymbol{A}$ 的秩等于 n; 再由定理 12, $\boldsymbol{A}$ 的秩等于 n 当且仅当 $\boldsymbol{A}$ 是非奇异的. 故推论 5 得证.

习题

1. 通过计算下列矩阵的行秩和列秩验证定理 19 的推论 3:
   (a) 7.6 节习题 1 中的矩阵.
   (b) 7.6 节习题 7(a) 和 7(b) 中的矩阵.
2. 对 7.6 节习题 2 的每个矩阵, 求出等价的对角矩阵.
3. 对 7.6 节习题 7 的矩阵, 求出等价的对角矩阵.
4. 设 T 是 m 维向量空间 V 到 n 维向量空间 W 的线性变换, 证明: 在 V 和 W 中适当地选取基, 使得 T 的方程取成形式 $y_i = x_i (i = 1, \cdots, r), y_j = 0 (j = r+1, \cdots, n)$.
5. (a) 证明: 任意初等矩阵的转置还是初等矩阵.
   (b) 用 (a) 证明: 非奇异矩阵的转置还是非奇异矩阵.

*6. 证明: 如果 $n \times n$ 矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的秩分别为 r 和 s, 那么 $\boldsymbol{AB}$ 的秩不少于 $(r+s)-n$. (提示: 利用 $\boldsymbol{A}$ 的标准型).

*7. (a) 证明零度的西尔维斯特 (Sylvester) 定律: 乘积 $\boldsymbol{AB}$ 的零度绝不超过这两个因子的零度之和, 并且绝不少于 $\boldsymbol{A}$ 的零度. 如果 $\boldsymbol{A}$ 是方阵, 则 $\boldsymbol{AB}$ 的零度至少等于 $\boldsymbol{B}$ 的零度.
   (b) 给出例子说明乘积 $\boldsymbol{AB}$ 的零度可以达到上述两种界限情况.

8. 证明: 对角线元素全不为零的任意 $n \times n$ 非奇异矩阵 $\boldsymbol{P}$ 可以写成 $\boldsymbol{P} = \boldsymbol{T}\boldsymbol{U}^{\mathrm{T}}$, 其中 $\boldsymbol{T}$ 和 $\boldsymbol{U}$ 都是三角形矩阵.
9. 证明: 一个 $m \times n$ 矩阵 $\boldsymbol{A}$ 的秩至多是 1 当且仅当它可以表示成乘积 $\boldsymbol{A} = \boldsymbol{BC}$, 其中 $\boldsymbol{B}$ 是 $m \times 1$ 矩阵, $\boldsymbol{C}$ 是 $1 \times n$ 矩阵.
10. 证明: 任意秩为 r 的矩阵等于 r 个秩为 1 的矩阵之和.

*11. 设一系列初等行运算 $\boldsymbol{E}_1, \cdots, \boldsymbol{E}_r$ 适当交叉一系列初等列运算 $\boldsymbol{E}'_1, \cdots, \boldsymbol{E}'_s$, 把矩阵 $\boldsymbol{A}$ 化为 $\boldsymbol{I}$. 证明: $\boldsymbol{A}^{-1} = \boldsymbol{QP}$, 这里 $\boldsymbol{P} = \boldsymbol{E}_r \cdots \boldsymbol{E}_1, \boldsymbol{Q} = \boldsymbol{E}'_1 \cdots \boldsymbol{E}'_s$ 是通过单位矩阵 $\boldsymbol{I}$ 进行一系列相同的初等运算而得到的矩阵.

12. 证明: 像定理 18 那样, 如果 $\boldsymbol{PQ} = \boldsymbol{D}$, 那么联立线性方程组 $\boldsymbol{AX}^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}$(8.8 节), 可以通过求解方程 $\boldsymbol{DY}^{\mathrm{T}} = \boldsymbol{PB}^{\mathrm{T}}$, 然后再计算 $\boldsymbol{X}^{\mathrm{T}} = \boldsymbol{QY}^{\mathrm{T}}$, 来求解 $\boldsymbol{X}^{\mathrm{T}}$.

*8.10 双线性函数与张量积

现在设 V 和 W 是同一域上的任意两个向量空间. 如果两个变量 $\boldsymbol{\xi} \in V$ 和

$\boldsymbol{\eta}\in W$ 的函数 $f(\boldsymbol{\xi},\boldsymbol{\eta})$ 在 F 中取值, 满足对所有 $\boldsymbol{\xi},\boldsymbol{\xi}'\in V$ 和所有 $\boldsymbol{\eta},\boldsymbol{\eta}'\in W$ 有

$$f(a\boldsymbol{\xi}+b\boldsymbol{\xi}',\boldsymbol{\eta})=af(\boldsymbol{\xi},\boldsymbol{\eta})+bf(\boldsymbol{\xi}',\boldsymbol{\eta}), \tag{58}$$

$$f(\boldsymbol{\xi},c\boldsymbol{\eta}+d\boldsymbol{\eta}')=cf(\boldsymbol{\xi},\boldsymbol{\eta})+df(\boldsymbol{\xi},\boldsymbol{\eta}'), \tag{58'}$$

那么称 $f(\boldsymbol{\xi},\boldsymbol{\eta})$ 是双线性函数. 重复在证明 7.12 节定理 23 时用过的论证方法, 我们容易得到下面的结果.

定理 20 如果 V 和 W 分别具有有限基 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_m$ 和 $\gamma_1,\cdots,\gamma_n$, 那么变量为 $\boldsymbol{\xi}=x_1\boldsymbol{\beta}_1+\cdots+x_m\boldsymbol{\beta}_m$ 和 $\boldsymbol{\eta}=y_1\gamma_1+\cdots+y_n\gamma_n$ 的双线性函数具有形式

$$f(\boldsymbol{\xi},\boldsymbol{\eta})=\sum_{i=1}^{m}\sum_{j=1}^{n}x_ia_{ij}y_j,\qquad a_{ij}=f(\boldsymbol{\beta}_i,\gamma_j). \tag{59}$$

注意, (59) 式的两个方程描述了域 F 上的矩阵 $\boldsymbol{A}=(a_{ij})$ 和双线性函数 $f: F^m\times F^n\to F$ 之间的可逆函数 $\boldsymbol{A}\longmapsto f$ 和 $f\longmapsto \boldsymbol{A}$, 这里 $F^m\times F^n$ 是 1.11 节中定义的 F^m 和 F^n 的笛卡儿积. 因此 (59) 表示的对应是双射.

上述双射能够推广. 我们可以定义一个双线性函数 $h(\boldsymbol{\xi},\boldsymbol{\eta})$, 它的变量 $\boldsymbol{\xi}$ 和 $\boldsymbol{\eta}$ 分别在向量空间 V 和 W 中, 函数值取在第三个向量空间 U 中 (U,V 和 W 是同一个域 F 上的向量空间). 也就是说, 这样的函数 $h:V\times W\to U$, 当它满足 (58) 和 (58′) 时, 就称为是双线性的.

存在很多这种函数. 例如, $\mathbf{R}^3$ 中两个向量的外积 $\boldsymbol{\xi}\times\boldsymbol{\eta}$ 是一个双线性函数, 其中取 $U=V=W=\mathbf{R}^3$. 同样, 如果我们设 $U=V=W=M_n$ 是域 F 上所有 $n\times n$ 矩阵组成的向量空间, 那么, 如定理 3 和定理 5 所述, "矩阵乘积" 函数 $p(\boldsymbol{A},\boldsymbol{B})=\boldsymbol{AB}$ 是从 $M_n\times M_n$ 到 M_n 的双线性函数.

上面定理 20 的结论对于前面说的更一般的情况也是成立的, 其证明类似.

定理 21 设 F 上的向量空间 V 和 W 分别具有有限基 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_m$ 和 $\gamma_1,\cdots,\gamma_n$. 那么, F 上第三个向量空间 U 中的任意 mn 个向量 θ_{ij} 确定一个双线性函数 $h: V\times W\to U$, 它由公式

$$h(\boldsymbol{\xi},\boldsymbol{\eta})=\sum_{i=1}^{m}\sum_{j=1}^{n}x_iy_j\theta_{ij} \tag{60}$$

给出, 其中 $\boldsymbol{\xi}\in V,\boldsymbol{\eta}\in W$. 而且, 任意双线性函数 $h:V\times W\to U$ 都可表示为 (60) 的形式, 其中 $\theta_{ij}=h(\boldsymbol{\beta}_i,\gamma_j)$, 所以 $H\longmapsto h$ 是从所有 $m\times n$ 矩阵 $H=(\theta_{ij})(\theta_{ij}\in U)$ 的集合到双线性函数 $h:V\times W\to U$ 的集合的一个双射.

这个定理暗示给我们一种得到标准的或 "最一般的" $V\times W$ 上双线性函数 $\otimes$ 的方法, 这里符号 $\otimes$ 能常写在自变量中间, 如 $\boldsymbol{\xi}\otimes\boldsymbol{\eta}=\otimes(\boldsymbol{\xi},\boldsymbol{\eta})$. 这个函数 $\otimes$ 在一个新的向量空间中取值, 这个空间记作 $V\otimes W$. 事实上, 我们来构造这个空间, 使

它有一组基由 mn 个向量 $\boldsymbol{\alpha}_{ij}(i=1,\cdots,m;\ j=1,\cdots,n)$ 组成, 这些 $\boldsymbol{\alpha}_{ij}$ 是 $\otimes$ 作用到 V 和 W 的基向量上而取的值, 即 $\boldsymbol{\alpha}_{ij}=\boldsymbol{\beta}_i\otimes\boldsymbol{\gamma}_j$. 这就意味着函数 $\otimes$ 可以定义为

$$(x_1\boldsymbol{\beta}_1+\cdots+x_m\boldsymbol{\beta}_m)\otimes(y_1\boldsymbol{\gamma}_1+\cdots+y_n\boldsymbol{\gamma}_n)=\sum_{i=1}^{m}\sum_{j=1}^{n}x_iy_j\boldsymbol{\alpha}_{ij}. \tag{61}$$

同 (60) 式一样, 这里只是用 $\boldsymbol{\alpha}_{ij}$ 代替 (60) 中的 θ_{ij}. 然而, 这个新空间 $V\otimes W$ 最好是用一个不依赖于 V 和 W 的基的选择的固有性质来描述, 如下所述.

定理 22　*对于域 F 上任意给定的有限维向量空间 V 和 W, 存在向量空间 $V\otimes W$ 和双线性函数*

$$\otimes:V\times W\to V\otimes W$$

具有如下性质: 对于 F 上任意向量空间 U 的任意双线性函数 $h:V\times W\to U$ 可以通过 $\otimes:V\times W\to V\otimes W$ 表示成

$$h(\boldsymbol{\xi},\boldsymbol{\eta})=(\boldsymbol{\xi}\otimes\boldsymbol{\eta})T,\qquad \boldsymbol{\xi}\in V,\boldsymbol{\eta}\in W,$$

其中 T 是唯一的线性函数 $T:V\otimes W\to U$.

证明　我们首先按上面方法构造 $\otimes$. 然后像 (60) 那样, 可以通过 mn 个向量 $\boldsymbol{\theta}_{ij}=h(\boldsymbol{\beta}_i,\boldsymbol{\gamma}_j)$ 来表示任意双线性函数 h. 现在由 (60) 和 (61) 两式的平行关系引导出一个线性变换 $T:V\otimes W\to U$, 当把它看作把 $V\otimes W$ 的每个基向量 $\boldsymbol{\alpha}_{ij}$ 变到 U 中的 $\boldsymbol{\theta}_{ij}$ 的变换时, T 是唯一确定的. 那么公式 (60) 变成

$$h(\boldsymbol{\xi},\boldsymbol{\eta})=\sum\sum x_iy_j(\boldsymbol{\alpha}_{ij}T)=(\sum\sum x_iy_j\boldsymbol{\alpha}_{ij})T=(\boldsymbol{\xi}\otimes\boldsymbol{\eta})T,$$

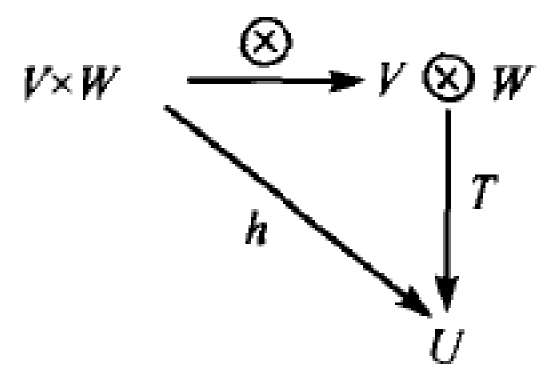
满足定理要求. 另一方面, 如果对某个线性变换 $T':V\otimes W\to U$, 有 $h(\boldsymbol{\xi},\boldsymbol{\eta})=(\boldsymbol{\xi}\otimes\boldsymbol{\eta})T'$, 那么有

$$\boldsymbol{\alpha}_{ij}T'=(\boldsymbol{\beta}_i\otimes\boldsymbol{\gamma}_j)T'=\theta_{ij},\quad i=1,\cdots,m;\quad j=1,\cdots,n,$$

于是 T' 一定是上面用过的 T. 因此 T 是唯一的, 如定理所述.

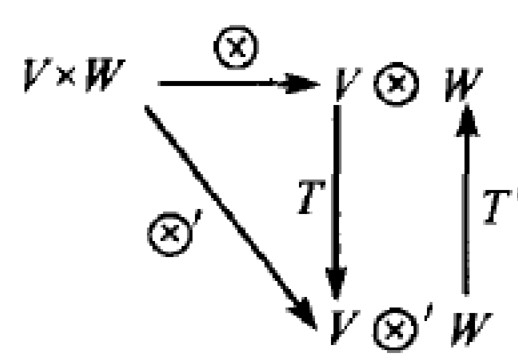
例　设 $V=F^m,W=F^n$, 设 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_m$ 和 $\gamma_1,\cdots,\gamma_n$ 分别是空间 V 和 W 的标准单位向量 $\varepsilon_1,\cdots,\varepsilon_m$ 和 $\varepsilon'_1,\cdots,\varepsilon'_n$. 那么 $V\otimes W=F^{mn}$ 可以是由所有 $m\times n$ 矩阵 (a_{ij}) 组成的空间, 而 $\otimes$ 把每个 $(\boldsymbol{\xi},\eta)\in V\times W$ 映射到秩为 1 的矩阵 $(x_iy_j)=(a_{ij})$. 每个双线性函数 $\theta:V\times W\to U$ 由 mn 个向量 $\theta(\varepsilon_i,\varepsilon'_j)=h_{ij}$ 来确定. 那么函数 θ 显然是 $\otimes$ 和线性函数 $T:V\otimes W\to U$ 的合成 $\otimes T$, $\otimes$ 像上面那样定义, T 用公式 $[(a_{ij})]T=\sum a_{ij}h_{ij}$ 来定义, 这因为对所有 $\boldsymbol{\xi}\in V,\boldsymbol{\eta}\in W$, 有 $(\boldsymbol{\xi}\otimes\boldsymbol{\eta})T=\sum x_iy_jh_{ij}$.

泛性质(Universality)　这个定理可用图表示如下



图中顶上一行是“标准”双线性函数 $\otimes$, 底下一行是任意双线性函数 h; 这个定理表明, 总是恰好存在一个线性变换 T, 使得图按照 $\otimes T = h$ “画出”, 即使得 $h(\boldsymbol{\xi},\boldsymbol{\eta}) = (\boldsymbol{\xi}\otimes\boldsymbol{\eta})T$. 由于这个原因, $\otimes$ 称为泛双线性函数, 而其他任何双线性函数 h 可以由它得到.

特别是, 如果我们构造另一个任意的标准双线性函数 $\otimes'$, 也具有同样的“泛性质”—— 比方说, 使用 V 和 W 的不同基 —— 我们将有图表示如下:



满足 $\otimes T = \otimes'$ 和 $\otimes' T' = \otimes$. 这就意味着 $\otimes TT' = \otimes = \otimes I$, 其中 I 是恒等变换. 根据定理, 这又意味着 $TT' = I$. 类似地有 $T'T = I$, 所以 T 是可逆变换, 它的逆是 T', 因此 T 是一个同构 $V\otimes W \cong V\otimes' W$.

具有“泛性质”的空间 $V\otimes W$ 称为空间 V 和 W 的张量积. 上段叙述的结论表明, 这“泛性质”唯一地 (在同构意义下) 确定这个空间. 例如, 我们不从基 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_m$ 和 $\gamma_1,\cdots,\gamma_n$ 来构造 $V\otimes W$, 而是从 V 和 W 的另外不同的基来构造, 得到一个同构空间 $V\otimes W$. 就这一点而言, 这张量积空间 $V\otimes W$ 可以不用任意基 (对于无穷维空间 V 和 W 用无穷多个基向量) 而用其他方法来构造, 它总是具有相同的“泛性质”. 我们特别用它的基 $\boldsymbol{\beta}_i\otimes\gamma_j$ 来构造, 就会看到, 它的维数是

$$\dim(V\otimes W) = \dim V + \dim W.$$

另外, 当给定一个空间 V 和它的对偶空间 V^*, 我们可以构造各种张量积:

$$V\otimes V,\ V\otimes V\otimes V,\cdots,V\otimes V^*,\ V\otimes V^*\otimes V,\cdots$$

在微分几何和相对论中用到这些张量空间.

习 题

1. 证明: 由 (59) 定义的映射 $f\longmapsto \boldsymbol{A}$ 是向量空间的一个同构, 它是从 $V\times W$ 上所有双线性函数的空间到 F 上所有 $m\times n$ 矩阵的空间的同构.

2. 证明：公式 $q(x) = a(x)p'(x)$ 定义了一个双线性函数 $\phi(a,p) = q$，它是从所有实多项式的空间做成的笛卡儿积 $\mathbf{R}[x] \times \mathbf{R}[x]$ 到 $\mathbf{R}[x]$ 的一个双线性函数.
3. 证明：函数 $p(\boldsymbol{A},\boldsymbol{B}) = \boldsymbol{AB}$ 是从 $V \times W$ 到 U 的双线性函数，这里 V 是 F 上所有 $m \times r$ 矩阵的空间，W 是 F 上所有 $r \times n$ 矩阵的空间. U 是什么空间?

在习题 4 和习题 5 中，设 U, V, W 是 F 上任意向量空间.

4. 建立下面的自然同构：

$$V \otimes F \cong V,\ V \otimes W \cong W \otimes V, U \otimes (V \otimes W) \cong (U \otimes V) \otimes W.$$

5. 证明：集合 $\mathrm{Hom}(V \otimes W, U) = \mathrm{Hom}(V, \mathrm{Hom}(W, U))$. ($\mathrm{Hom}(S,T)$) 的定义见 8.2 节末尾.)

*6. 在 $V \otimes W$ 中每个向量是元素 $\boldsymbol{\xi} \otimes \boldsymbol{\eta}$ 的和. 证明：存在不能表示成单个元素 $\boldsymbol{\xi} \otimes \boldsymbol{\eta}$ 的向量. (提示：取 $V = F^2 = W$.)

*7. $m \times m$ 矩阵 $\boldsymbol{A}$ 和 $n \times n$ 矩阵 $\boldsymbol{B}$ 的克罗内克尔积 $\boldsymbol{A} \otimes \boldsymbol{B}$ 是一个矩阵 $\boldsymbol{C}$，它的元素为 $c_{pq} = a_{ik}b_{jl}$，这里 p 和 q 是按适当次序排列的数对 (i,j) 和 $(k,\ l)$. $\boldsymbol{A} \otimes \boldsymbol{B}$ 同 $\boldsymbol{V} \otimes \boldsymbol{W}$ 上什么样的线性变换自然对应?

*8.11 四 元 数

对于方阵成立的代数定律可应用到其他代数系统中，例如哈密顿四元数，这些四元数组成实数域上的四维向量空间，它的基由四个特殊的向量构成，它们分别记作 $1, i, j, k$. 四元数的代数运算就是通常的两种向量运算 (向量加法和数乘运算)，再加上四元数乘法的新运算.

定义 *四元数是向量 $x = x_0 + x_1 i + x_2 j + x_3 k$，其中系数 x_0, x_1, x_2, x_3 为实数. 四元数 $1, i, j, k$ 中任意两个的乘积定义如下：首先要求把 1 当作单位元素，其他按照下表*

$$i^2 = j^2 = k^2 = -1$$

$$ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j. \tag{62}$$

如果 c 和 d 是任意标量，而 l, m，是 $1, i, j, k$ 中任意两个，那么乘积 $(cl)(dm)$ 定义为 $(cd)(lm)$. 这些法则连同分配律一起就确定了任意两个四元数的乘积.

例如，如果 $x = x_0 + x_1 i + x_2 j + x_3 k$ 和 $y = y_0 + y_1 i + y_2 j + y_3 k$ 是任意两个四元数，那么它们的乘积是

$$\begin{aligned} xy = {} & x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3 \\ & + (x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)i \\ & + (x_0y_2 + x_2y_0 + x_3y_1 - x_1y_3)j \\ & + (x_0y_3 + x_3y_0 + x_1y_2 - x_2y_1)k. \end{aligned} \tag{63}$$

虽然四元数乘法不满足交换律, 但是它们满足关于域的其他每个公设. 具有上述性质的数系称为体 (或可除环).

定义 满足下列条件的元素系统 R 称为体, 它在单值二元运算 —— 加法和乘法之下是封闭的, 并且

(i) 在加法之下, R 是含有单位元素 0 的交换群;

(ii) 在乘法之下, 全体不为 0 的元素构成群;

(iii) 两种分配律都成立

$$a(b+c)=ab+ac \quad \text{和} \quad (a+b)c=ac+bc.$$

从这些公设容易导出法则 $a0=0a=0$, 因而导出含有因子 0 的乘法结合律. 由此推出, 任意可交换的体是一个域. 我们还看到, 在体上与 8.1 节 ~8.7 节的结果相类似的结果也成立, 只要我们注意一下出现标量因子的那一边. 例如, 对于向量 $\boldsymbol{\xi}$ 与标量 c 的乘积 $c\boldsymbol{\xi}$, 我们把标量 c 写在左边, 而在定义变换 T 和标量 c 的乘积时 (8.2 节), 我们把标量写在右边 $\boldsymbol{\xi}(Tc)=(c\boldsymbol{\xi})T$, 同样, 矩阵与标量相乘时, 我们把标量写在右边. 于是体 R 上左向量空间的线性变换 T 组成的空间是 R 上的右向量空间.

定理 23 全体四元数构成一个体.

每个公设的证明, 除了乘法逆的存在性 (根据 6.4 节定理 3, 它可推出消去律) 和乘法结合律以外, 都是显然的. 为了证明每个非零四元素 $x=x_0+x_1i+x_2j+x_3k$ 有逆, 我们定义 x 的共轭数 $x^*=x_0-x_1i-x_2j-x_3k$. 那么容易看出, 用 $N(x)=xx^*$ 定义的 x 的范数是一个实数, 它满足

$$N(x)=xx^*=x^*x=x_0^2+x_1^2+x_2^2+x_3^2>0, \quad \text{当} x\neq 0. \tag{64}$$

因此 x 的逆是 $\dfrac{x^*}{N(x)}$.

利用复数很容易完成结合律的证明. 诚然, 从 (64) 式容易看出, $x_2=x_3=0$ 的全体四元数 $x=x_0+x_1i$ 构成一个与复数域同构的子系统. 此外, 有

$$x=(x_0+x_1i)+(x_2+x_3i)j=z_1+z_2j, \tag{65}$$

其中 z_1 和 z_2 的性质很像普通的复数. 实际上, (62) 的所有运算法则包含在展开式 (65) 及结合律、分配律和法则

$$z_1j=jz_1^*, \quad j^2=-1 \tag{66}$$

中, 式中 $z_1^*=x_0-x_1i$ 是 $z_1=x_0+x_1i$ 的复共轭 (并且是四元数共轭!). 诚然, 两个形为 (65) 的四元数的乘积是

$$(z_1+z_2j)(w_1+w_2j)=(z_1w_1-z_1w_2^*)+(z_1w_2+z_2w_1^*)j.$$

用这个公式, 我们可以很容易地验证结合律.

每个四元数 x 满足一个以 x 和 x^* 为根的实系数二次方程 $f(t)=0$. 这个方程是

$$f(t)=(t-x)(t-x^*)=t^2-(x+x^*)t+xx^*=t^2-2x_0t+N(x).$$

任意四元数 $x=x_0+x_1i+x_2j+x_3k$ 可以分解为它的实数部分 x_0 和它的"纯四元数"部分 $x_1i+x_2j+x_3k$. 它们有各种有趣的性质 (参见习题 2(c) 和习题 15), 最稀奇的一个性质是关于纯四元数 $\boldsymbol{\xi}=x_1i+x_2j+x_3k$ 和 $\boldsymbol{\eta}=y_1i+y_2j+y_3k$ 的乘法. 根据定义, 有

$$\boldsymbol{\xi\eta}=\boldsymbol{\xi}\times\boldsymbol{\eta}-(\boldsymbol{\xi},\boldsymbol{\eta}), \tag{67}$$

式中 $\boldsymbol{\xi}\times\boldsymbol{\eta}=(x_2y_3-x_3y_2)i+(x_3y_1-x_1y_3)j+(x_1y_2-x_2y_1)k$ 是通常的 $\boldsymbol{\xi}$ 和 $\boldsymbol{\eta}$ 的外积 (即向量积), $(\boldsymbol{\xi},\boldsymbol{\eta})=x_1y_1+x_2y_2+x_3y_3$ 是第 7 章中定义过的内积. 就是由于这个恒等式 (67), 从 1850 年到 1900 年这半世纪里, 很多近代三维向量空间分析都用四元数的语言来表达.

1944 年艾兰伯格 (Eilenberg) 和尼文 (Niven) 证明了, 任何四元数系数多项式方程 $f(x)=a_0+a_1x+\cdots+a_nx^n=0$(其中 $a_n\neq 0, n>0$) 具有一个四元数解.

习 题

1. 分别对下列两种情况解方程 $xc=d$:

   (a) $c=i,\ d=1+j$; (b) $c=2+j, d=3+k$.

2. (a) 证明: $x^2=-1$ 有无穷多个四元数解 x.

   (b) 说明这同 3.2 节关于多项式根数的定理 3 的推论为什么不矛盾.

   (c) 证明: 实四元数是其平方后为正实数的那些四元数, 而纯四元数是其平方后为负实数的那些四元数. 并证明: 满足条件 $x^2<0$ 的四元数组成的集合在加法和减法之下是封闭的.

   (d) 证明: 如果 q 不是实数, 那么 $x^2=q$ 恰有两个四元数解.

3. 设 $a=1+i+j,\ b=1+j+k$,

   (a) 求 $a+b,\ ab,\ a-b,\ ia-2b,\ a^*,\ aa^*$.

   (b) 解方程 $ax=b, xa=b, x^2=b, bx+(2j+k)=a$.

4. 从 (62) 式导出乘法表 (66).

5. (a) 证明: x 的范数 $N(x)=xx^*$ 是 $x_0^2+x_1^2+x_2^2+x_3^2$.

   (b) 证明: $x^*y^*=(yx)^*$.

6. 证明: 在非零四元数组成的乘法群中, 它的中心确实是由全体非零实四元数组成.

7. 证明: 当 $a\neq 0$ 时, 四元数方程 $xa=b$ 的解是唯一确定的.

8. 证明: 如果四元数 x 满足含有实系数 a_0 和 b_0 的二次方程 $x^2+a_0x+b_0=0$, 那么每个四元数 $q^{-1}xq(q\neq 0)$ 满足相同的二次方程.

9. 证明：四元数乘法满足结合律. (提示：用 (65) 式和 (66) 式.)
10. 在四元数代数中证明：元素 $\pm 1, \pm i, \pm j, \pm k$ 构成一个乘法群. (这个群可以直接定义, 它称为四元数群.)
11. (a) 列举四元数群 (习题 10) 的全体子群, 并证明它们都是正规子群.
    (b) 证明：四元数群与正方形对称群不同构.
12. (a) 证明：全体具有有理系数 x_i 的四元数 $x = x_0 + x_1 i + x_2 j + x_3 k$ 构成一个体.
    (b) 证明：对于具有复系数的四元数, 情况不是这样. (注意：不要把 $\sqrt{-1} \in \mathbf{C}$ 与四元数单位 i 混同起来.)
13. 证明：在体中, 加法交换律可从其他公设推出. (提示：按两种不同的方式展开 $(a+b)(1+1)$.)
14. 如果把 $\dfrac{a}{b}$ 解释为 ab^{-1}, 那么你能够证明 2.1 节定理 2 的多少个条件在一般体中成立?
15. 证明：两个向量的外积不满足结合律.
16. 证明：如果两个整数 a 和 b 都是四个整数的平方和, 那么乘积 ab 也是四个整数的平方和. (提示：用习题 5.)
17. 从 $i^2 = j^2 = k^2 = ijk = -1$, 推导 (62) 的所有运算法则.
18. 对于以四元数为元素的矩阵, 公式 $(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$ 成立吗?

第9章 线 性 群

9.1 基的变换

在向量空间 V 中, 向量 $\boldsymbol{\xi}$ 的坐标依赖于 V 的基的选取 (见 7.8 节), 因此, 基的任意变化将引起 $\boldsymbol{\xi}$ 的坐标的变化. 例如, 在实平面 $\mathbf{R}^2$ 中, 向量 $\boldsymbol{\beta} = 4\boldsymbol{\varepsilon}_1 + 2\boldsymbol{\varepsilon}_2$, 按照定义, 它关于由单位向量 $\boldsymbol{\varepsilon}_1$ 和 $\boldsymbol{\varepsilon}_2$ 组成的基的坐标是 (4, 2). 向量

$$\boldsymbol{\alpha}_1 = 2\boldsymbol{\varepsilon}_1,\ \boldsymbol{\alpha}_2 = \boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2 \tag{1}$$

也可以构成一组基, 关于这组基, β 可表示成 $\boldsymbol{\beta} = \boldsymbol{\alpha}_1 + 2\boldsymbol{\alpha}_2$. 系数 1 和系数 2 是 β 关于这组新基的坐标 (也就是关于图 9-1 所表示的斜角坐标系的坐标).

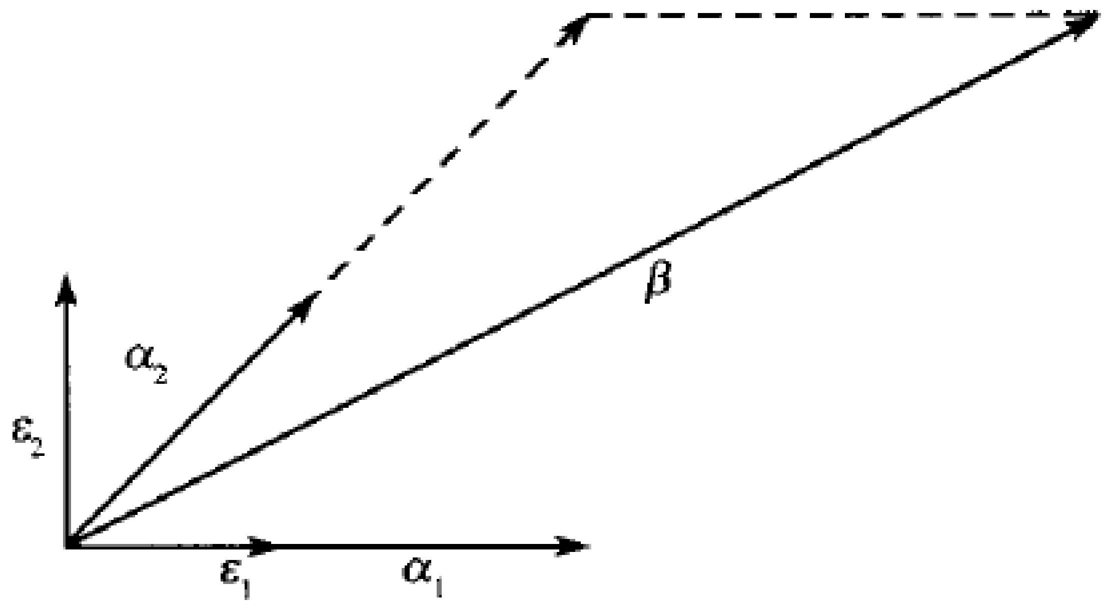


图 9-1

更一般地, 任意向量 ξ 关于新基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 的坐标 x_1^*, x_2^* 可以从 $\boldsymbol{\xi}$ 的 "老" 坐标 x_1, x_2 按下面方法求得. 按照定义 (7.8 节), 这两组坐标是向量 $\boldsymbol{\xi}$ 关于两组基表达式的系数

$$\boldsymbol{\xi} = x_1\boldsymbol{\varepsilon}_1 + x_2\boldsymbol{\varepsilon}_2, \qquad \boldsymbol{\xi} = x_1^*\boldsymbol{\alpha}_1 + x_2^*\boldsymbol{\alpha}_2.$$

解向量方程 (1), 我们求出 $\boldsymbol{\varepsilon}_1$ 和 $\boldsymbol{\varepsilon}_2$:

$$\boldsymbol{\varepsilon}_1 = \frac{1}{2}\boldsymbol{\alpha}_1, \qquad \boldsymbol{\varepsilon}_2 = \boldsymbol{\alpha}_2 - \frac{1}{2}\boldsymbol{\alpha}_1.$$

把 $\boldsymbol{\varepsilon}_1$ 和 $\boldsymbol{\varepsilon}_2$ 的值代入 $\boldsymbol{\xi}$ 的第一个表达式中, 我们得到

$$\boldsymbol{\xi} = x_1\left(\frac{1}{2}\boldsymbol{\alpha}_1\right) + x_2\left(\boldsymbol{\alpha}_2 - \frac{1}{2}\boldsymbol{\alpha}_1\right) = \frac{1}{2}(x_1 - x_2)\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2.$$

因此 $\boldsymbol{\xi}$ 的新坐标由线性齐次方程

$$x_1^* = \frac{1}{2}(x_1 - x_2), \qquad x_2^* = x_2 \tag{2}$$

给出. 反过来, 老坐标可以通过新坐标表示为

$$x_1 = 2x_1^* + x_2^*, \qquad x_2 = x_2^*.$$

在 n 维空间中有类似的关系式. 如果 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 是一组给定的基, 这些向量是按一定次序排列的, 而 $\boldsymbol{\alpha}_1^*, \cdots, \boldsymbol{\alpha}_n^*$ 是一组新的 (有序) 基, 那么新基的每个向量 $\boldsymbol{\alpha}_i^*$ 可以表示为老基向量的线性组合:

$$\boldsymbol{\alpha}_i^* = p_{i1}\boldsymbol{\alpha}_1 + \cdots + p_{in}\boldsymbol{\alpha}_n = \sum_{j=1}^{n} p_{ij}\boldsymbol{\alpha}_j,\ i = 1, \cdots, n. \tag{3}$$

表达式 (3) 可以形式地写成矩阵方程 $\boldsymbol{\alpha}^* = \boldsymbol{\alpha}\boldsymbol{P}^{\mathrm{T}}$, 这里 $\boldsymbol{P}^{\mathrm{T}}$ 是 $\boldsymbol{P}$ 的转置矩阵, $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n)$, $\boldsymbol{\alpha}^* = (\boldsymbol{\alpha}_1^*, \cdots, \boldsymbol{\alpha}_n^*)$.

这些表达式的系数矩阵 $\boldsymbol{P} = (p_{ij})$ 的第 i 行元素是向量 $\boldsymbol{\alpha}_i^*$ 的老坐标 $(p_{i1}, \cdots, p_{in})$. 因为向量 $\boldsymbol{\alpha}_1^*, \cdots, \boldsymbol{\alpha}_n^*$ 构成一组基, 所以 $\boldsymbol{P}$ 的所有行是线性无关的, 因此 $\boldsymbol{P}$ 是非奇异的 (8.6 节定理 9). 反过来, 如果 $\boldsymbol{P} = (p_{ij})$ 是任意非奇异矩阵, $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 是 V 的任意一组基, 那么用矩阵 $\boldsymbol{P}$ 按公式 (3) 所确定的向量 $\boldsymbol{\alpha}_1^*, \cdots, \boldsymbol{\alpha}_n^*$ 是线性无关的, 因此构成 V 的一组新基. 这就证明了

定理 1 如果 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 是向量空间 V 的一组基, 那么对每个非奇异矩阵 $\boldsymbol{P} = (p_{ij})$, n 个向量 $\boldsymbol{\alpha}_i^* = \Sigma p_{ij}\boldsymbol{\alpha}_j (i = 1, \cdots, n)$ 构成 V 的一组新基, 并且 V 的每组基可以按这种方法恰由一个非奇异矩阵 $\boldsymbol{P}$ 得到.

我们还可以用新基表示老基, 其方程式是 $\boldsymbol{\alpha}_k = \Sigma q_{ki}\boldsymbol{\alpha}_i^*$, 其中系数矩阵为 $\boldsymbol{Q} = (q_{ki})$. 如果把用 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 表示的 $\boldsymbol{\alpha}_i^*$ 的值代入这个方程式中, 我们就得到

$$\boldsymbol{\alpha}_k = \sum_i q_{ki}\Big(\sum_j p_{ij}\boldsymbol{\alpha}_j\Big) = \sum_j \Big(\sum_i q_{ki}p_{ij}\Big)\boldsymbol{\alpha}_j.$$

可是, $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_k$ 这组向量用它们本身来表示的话, 只有一种表达式, 即 $\boldsymbol{\alpha}_k = \boldsymbol{\alpha}_k$. 因此这里 $\boldsymbol{\alpha}_j$ 的系数 $\sum q_{ki}p_{ij}$ 一定是 0 或 1, 它取决于 $k \neq j$ 或 $k = j$. 由于这些系数正好是乘积矩阵 $\boldsymbol{QP}$ 的 (k, j) 位置上的元素, 因此 $\boldsymbol{QP} = \boldsymbol{I}$, 所以 $\boldsymbol{Q} = \boldsymbol{P}^{-1}$ 是 $\boldsymbol{P}$ 的逆矩阵.

与此平行的关于坐标变换的结果如下所述:

定理 2 如果向量空间 V 的基 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 变换成一组新基 $\boldsymbol{\alpha}_1^*, \cdots, \boldsymbol{\alpha}_n^*, \boldsymbol{\alpha}_i^*$ 表示成形式 $\boldsymbol{\alpha}_i^* = \sum_j p_{ij}\boldsymbol{\alpha}_j$, 那么任意向量 $\boldsymbol{\xi}$ 关于老基 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 的坐标 $x_1, \cdots, x_n$ 通过齐次线性方程组

$$x_j = x_1^* p_{1j} + \cdots + x_n^* p_{nj} = \sum_{i=1}^{n} x_i^* p_{ij} \tag{4}$$

可以确定 $\boldsymbol{\xi}$ 关于基 $\boldsymbol{\alpha}_1^*,\cdots,\boldsymbol{\alpha}_n^*$ 的新坐标 $x_1^*,\cdots,x_n^*$.

证明 按照定义 (7.8 节), $\boldsymbol{\xi}$ 关于基 $\boldsymbol{\alpha}_1^*,\cdots,\boldsymbol{\alpha}_n^*$ 的坐标 $x_1^*,\cdots,x_n^*$, 是把 $\boldsymbol{\xi}$ 看作 $\boldsymbol{\alpha}_1^*,\cdots,\boldsymbol{\alpha}_n^*$ 的线性组合 $\boldsymbol{\xi}=\Sigma x_i^*\boldsymbol{\alpha}_i^*$ 时, 表达式中的系数. 把关于 $\boldsymbol{\alpha}_i^*$ 的公式 (3) 代入这个表达式中, 便得出

$$\boldsymbol{\xi}=\sum_i x_i^*\Big(\sum_j p_{ij}\boldsymbol{\alpha}_j\Big)=\sum_j\Big(\sum_i x_i^*p_{ij}\Big)\boldsymbol{\alpha}_j.$$

这里每个 $\boldsymbol{\alpha}_j$ 的系数是 $\boldsymbol{\xi}$ 的老坐标 x_j, 因此方程组 (4) 成立.

方程组 (4) 还可以写成矩阵形式 $\boldsymbol{X}=\boldsymbol{X}^*\boldsymbol{P}$, 这里 $\boldsymbol{X}=(x_1,\cdots,x_n)$ 是老坐标的行矩阵, $\boldsymbol{X}^*=(x_1^*,\cdots x_n^*)$ 是新坐标的行矩阵. 因为 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 和 $\boldsymbol{\alpha}_1^*,\cdots,\boldsymbol{\alpha}_n^*$ 都是基, 所以 $\boldsymbol{P}$ 是非奇异的, 并且可以利用 $\boldsymbol{X}$ 表示 $\boldsymbol{X}^*$ 为 $\boldsymbol{X}^*=\boldsymbol{X}\boldsymbol{P}^{-1}$.

如果我们把这个矩阵方程同前面已提到的 (3) 式的矩阵公式 $\boldsymbol{\alpha}^*=\boldsymbol{\alpha}\boldsymbol{P}^{\mathrm{T}}$ 进行比较, 便可得到有趣的关系

$$\text{基:}\ \boldsymbol{\alpha}^*=\boldsymbol{\alpha}\boldsymbol{P}^{\mathrm{T}},\ \text{坐标:}\ \boldsymbol{X}^*=\boldsymbol{X}\boldsymbol{P}^{-1}. \tag{5}$$

第二个方程中的矩阵 $\boldsymbol{P}^{-1}$ 是第一个方程中的矩阵 $\boldsymbol{P}^{\mathrm{T}}$ 的*转置逆矩阵*. (有时把这些叙述概括成：坐标变换是基变换的转置逆变换.)

习 题

1. 设 T 是把通常的单位向量 ε_i(在 V_2 或 V_3 中) 变换到下面指定的向量 α_i. 求出相应的用老坐标表示新坐标的方程, 和用新坐标表示老坐标的方程. 对情况 (a) 和 (b) 画出图来.
   (a) $\boldsymbol{\alpha}_1=(1,1),\ \boldsymbol{\alpha}_2=(1,-1)$.
   (b) $\boldsymbol{\alpha}_1=(2,3),\ \boldsymbol{\alpha}_2=(-2,-1)$.
   (c) $\boldsymbol{\alpha}_1=(1,1,0),\ \boldsymbol{\alpha}_2=(1,0,1),\ \boldsymbol{\alpha}_3=(0,1,1)$.
   (d) $\boldsymbol{\alpha}_1=(\mathrm{i},1,\mathrm{i}),\ \boldsymbol{\alpha}_2=(0,1,\mathrm{i}),\ \boldsymbol{\alpha}_3=(0,\mathrm{i},1)$, 这里 $\mathrm{i}^2=-1$.
2. 如果新基 $\boldsymbol{\alpha}_1^*,\cdots,\boldsymbol{\alpha}_n^*$ 是通过形为 $\boldsymbol{\alpha}_i=\sum_j q_{ij}\boldsymbol{\alpha}_j^*(i=1,\cdots,n)$ 的方程组间接给出, 算出相应的坐标变换的方程.
3. 给出平面上坐标轴旋转 θ 角的坐标变换的方程.

9.2 相似矩阵与特征向量

向量空间 V 的线性变换 $T:V\to V$, 可以用各种不同的矩阵来表示, 这些矩阵依赖于 V 的基 (坐标系) 的选择. 比如, 在平面上, 由 $\varepsilon_1\longmapsto 3\varepsilon_1,\varepsilon_2\longmapsto -\varepsilon_1+2\varepsilon_2$

定义的变换, 它在 $\mathbf{R}^2$ 的普通坐标系中可用矩阵 $\boldsymbol{A}$ 表示出来, 矩阵 $\boldsymbol{A}$ 的行是 ε_1 和 ε_2 的变换式的坐标, 如下所示:

$$\boldsymbol{A}=\begin{pmatrix} 3 & 0 \\ -1 & 2 \end{pmatrix}, \quad \boldsymbol{D}=\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$$

但是对于 9.1 节中所讨论的新基 $\boldsymbol{\alpha}_1 = 2\varepsilon_1, \boldsymbol{\alpha}_2 = \varepsilon_1 + \varepsilon_2$, 上述变换便是 $\boldsymbol{\alpha}_1 \longmapsto 3\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \longmapsto 2\boldsymbol{\alpha}_2$; 因此它就可以用比较简单的对角矩阵来表示. 我们称这样两个矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{D}$ 是相似的.

为了推广这个结果, 让我们回想一下, 矩阵是怎样表示变换的. 取向量空间 V 的任意一组 (有序) 基 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 和任意线性变换 $T: V \to V$. 那么基向量 $\boldsymbol{\alpha}_i$ 在 T 之下的像可以用 8.1 节公式 (9) 写成

$$\alpha_i T = \sum_j a_{ij}\alpha_j, \quad \boldsymbol{A} = (a_{ij}). \tag{6}$$

因此, 对于基 $\boldsymbol{\alpha} = \{\boldsymbol{\alpha}_1, \cdots, \alpha_n\}$, T 用 $n \times n$ 矩阵 $\boldsymbol{A}$ 表示. 这个关系也可以通过坐标来表示. 设 $\boldsymbol{\xi} = \Sigma x_i \boldsymbol{\alpha}_i$ 是 V 的一个向量, 它对于基 $\boldsymbol{\alpha}$ 的坐标是 n- 数组 $\boldsymbol{X} = (x_1, \cdots, x_n)$. 那么像 $\boldsymbol{\eta} = \boldsymbol{\xi} T$ 是

$$\boldsymbol{\xi} T = (\Sigma x_i \boldsymbol{\alpha}_i)T = \Sigma x_i(\boldsymbol{\alpha}_i T) = \sum_i \sum_j x_i a_{ij}\alpha_j = \sum_j \Big(\sum_i x_i a_{ij}\Big)\boldsymbol{\alpha}_j.$$

而 $\boldsymbol{\eta}$ 的坐标 y_j 恰好是 $\boldsymbol{\alpha}_j$ 的系数, 所以

$$y_j = \sum_i x_i a_{ij},$$

$\boldsymbol{\eta}$ 的坐标向量 $\boldsymbol{Y}$ 恰好是矩阵乘积 $\boldsymbol{Y} = \boldsymbol{XA}$. 简明地写出就是

$$\boldsymbol{Y} = \boldsymbol{XA}, \text{其中 } \boldsymbol{X} \text{ 是 } \boldsymbol{\xi} \text{ 关于 } \boldsymbol{\alpha} \text{ 的坐标}, \boldsymbol{Y} \text{ 是 } \boldsymbol{\eta} = \boldsymbol{\xi} T \text{ 关于 } \boldsymbol{\alpha} \text{ 的坐标}. \tag{7}$$

两个等价的命题 (6) 和 (7) 都意味着, 对于基 $\boldsymbol{\alpha}$, 变换 T 可用矩阵 $\boldsymbol{A}$ 来表示.

现在设 $\boldsymbol{\alpha}_1^*, \cdots, \boldsymbol{\alpha}_n^*$ 是另一组基. 那么根据定理 1, 新基可以用一个 $n \times n$ 非奇异矩阵 $\boldsymbol{P}$ 通过老基来表示, 如 (3) 所示; 再根据定理 2, $\boldsymbol{\xi}$ 和 $\boldsymbol{\xi} T$ 的新坐标可以通过老坐标给出, 表示成 $\boldsymbol{X}^* = \boldsymbol{XP}^{-1}$ 和 $\boldsymbol{Y}^* = \boldsymbol{YP}^{-1}$. 那么由 (7) 有

$$\boldsymbol{Y}^* = \boldsymbol{YP}^{-1} = \boldsymbol{XAP}^{-1} = \boldsymbol{X}^*(\boldsymbol{PAP}^{-1}).$$

因此再根据 (7), 在新坐标系下表示变换 T 的矩阵 $\boldsymbol{B}$ 具有形式 $\boldsymbol{PAP}^{-1}$. 等价关系 $\boldsymbol{B} = \boldsymbol{PAP}^{-1}$ 在形式上很像群中的共轭元素的关系 (6.12 节). 在矩阵代数中, 这是很重要的, 称它为相似关系.

定义 元素在域 F 中的两个 $n\times n$ 矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$(在 F 上) 相似当且仅当在 F 上有一个 $n\times n$ 非奇异矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{B}=\boldsymbol{PAP}^{-1}$.

上述讨论就证明了

定理 3 域 F 上的两个 $n\times n$ 矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 对于 (通常) 不同的坐标系表示 F 上 n 维向量空间 V 的同一个线性变换 $T:V\to V$, 当且仅当矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是相似的.

我们还可以更清楚地把这个定理重述如下：

定理 3′ 假设对于 V 的基 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$, 线性变换 $T:V\to V$ 用矩阵 $\boldsymbol{A}$ 来表示, 设 $\boldsymbol{P}=(p_{ij})$ 是非奇异矩阵, $\boldsymbol{\alpha}_i^*=\sum\limits_j p_{ij}\boldsymbol{\alpha}_j$ 是 V 的相应的新基, 那么对于新基, T 就用矩阵 $\boldsymbol{PAP}^{-1}$ 来表示.

对角矩阵的代数运算特别容易：任意两个对角矩阵相加 (或相乘), 只是把相应的对角线元素相加 (或相乘). 由于这个以及其他理由, 考查什么样的矩阵同对角矩阵相似, 考查哪些成对的对角矩阵彼此是相似的, 这是非常重要的. 这些问题的回答包含了特征向量和特征根的概念, 这两个概念也称为本征向量和本征值.

定义 线性变换 $T:V\to V$ 的特征向量是 V 中满足条件 $\boldsymbol{\xi}T=c\boldsymbol{\xi}$ 的一个非零向量 $\boldsymbol{\xi}$, 这里 c 是某一标量; T 的特征值是满足 $\boldsymbol{\xi}T=c\boldsymbol{\xi}$ 的标量 c, 这里 $\boldsymbol{\xi}$ 是某一非零向量. 相应地, 方阵 $\boldsymbol{A}$ 的特征向量和特征值是满足 $\boldsymbol{XA}=c\boldsymbol{X}$ 的向量 $\boldsymbol{X}=(x_1,\cdots,x_n)$ 和标量 c. T(或 T_A) 的所有特征值的集合称为 T 的谱.

这样, T 的每个特征向量 $\boldsymbol{\xi}$ 确定一个特征值c, 并且每个特征值至少属于一个特征向量. 因为相似矩阵对应着不同基下的同一个线性变换, 所以相似矩阵具有同样的特征值. 显然, 如果向量 $\boldsymbol{X}\neq 0$ 对某一标量 c 满足 $\boldsymbol{XA}=c\boldsymbol{X}$, n 维向量 $\boldsymbol{X}$ 就是 $n\times n$ 矩阵 $\boldsymbol{A}$ 的特征向量. 如果矩阵 $\boldsymbol{B}=\boldsymbol{PAP}^{-1}$ 相似于 $\boldsymbol{A}$, 那么 $(\boldsymbol{XP}^{-1})\boldsymbol{B}=\boldsymbol{XP}^{-1}\boldsymbol{PAP}^{-1}=c(\boldsymbol{XP}^{-1})$, 所以 $\boldsymbol{XP}^{-1}$ 是 $\boldsymbol{B}$ 的属于同一特征值 c 的特征向量. 还应注意, 特征向量乘上任意非零标量还是特征向量.

特征向量与对角矩阵之间的联系由下面定理给出.

定理 4 一个 $n\times n$ 矩阵 $\boldsymbol{A}$ 与一个对角矩阵 $\boldsymbol{D}$ 相似当且仅当 $\boldsymbol{A}$ 的特征向量张成 F^n; 如果 $\boldsymbol{A}$ 与 $\boldsymbol{D}$ 相似, 那么 $\boldsymbol{A}$ 的特征值就是 $\boldsymbol{D}$ 的对角线元素.

特别是, 这个定理意味着, 对角矩阵的特征值是对角线上的元素.

证明 首先假定矩阵 $\boldsymbol{A}$ 与对角矩阵 $\boldsymbol{D}$ 相似, $\boldsymbol{D}$ 的对角线元素是 $d_1,\cdots,d_n$. 那么单位向量 $\boldsymbol{\varepsilon}_1=(1,0,\cdots,0),\cdots,\boldsymbol{\varepsilon}_n=(0,\cdots,0,1)$ 是 $\boldsymbol{D}$ 的特征向量, 这是因为 $\boldsymbol{\varepsilon}_1\boldsymbol{D}=d_1\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n\boldsymbol{D}=d_n\boldsymbol{\varepsilon}_n$. 还有, 对角线元素 $d_1,\cdots,d_n$ 是 $\boldsymbol{D}$ 的相应的特征值, 因此也是 $\boldsymbol{A}$ 的特征值. $d_1,\cdots,d_n$ 是唯一的一组特征值, 因为设 $\boldsymbol{X}=(x_1,\cdots,x_n)\neq 0$ 是 $\boldsymbol{D}$ 的任意特征向量, 那么对某一适当的特征值 c 有 $\boldsymbol{XD}=c\boldsymbol{X}$. 而 $\boldsymbol{XD}=(d_1x_1,\cdots,d_nx_n)$, 所以对所有的 i 有 $d_ix_i=cx_i$. 因为有某个 $x_i\neq 0$, 所以就证明了对这个 i, 有 $d_i=c$, 于是特征值 c 确实是某个 d_i.

反过来，假设 $\boldsymbol{A}$ 存在足够的特征向量张成整个空间 F^n，$T_{\boldsymbol{A}}$ 是 F^n 上相应的线性变换，那么 (7.4 节定理 4 的推论 2)，我们可以取出特征向量的一个子集合 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_n$，它构成 F^n 的一组基. 因为每个 $\boldsymbol{\beta}_i$ 是特征向量，所以有 $\boldsymbol{\beta}_1 T_{\boldsymbol{A}} = c_1\beta_1, \cdots, \beta_n T_{\boldsymbol{A}} = c_n\boldsymbol{\beta}_n$，其中 $c_1, \cdots, c_n$ 是一组特征值. 因此，对于基 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_n, T_{\boldsymbol{A}}$ 可以用对角矩阵 $\boldsymbol{D}$ 像公式 (6) 那样表示，这里 $\boldsymbol{D}$ 的对角线元素是 $c_1, \cdots, c_n$，所以 $\boldsymbol{A}$ 与这个矩阵 $\boldsymbol{D}$ 相似.

推论 如果矩阵 $\boldsymbol{P}$ 的行是 $n \times n$ 矩阵 $\boldsymbol{A}$ 的 n 个线性无关的特征向量，那么 $\boldsymbol{P}$ 是非奇异的，并且 $\boldsymbol{PAP}^{-1}$ 是对角矩阵.

证明 我们给出 n 个线性无关的 n 维向量 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$，它们是 $\boldsymbol{A}$ 的特征向量，所以对特征值 $c_1, \cdots, c_n$，有 $\boldsymbol{X}_i\boldsymbol{A} = c_i\boldsymbol{X}_i$, $i = 1, \cdots, n$. 以 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$ 为行的矩阵 $\boldsymbol{P}$ 是非奇异的，因为它所有的行是线性无关的. 根据分块矩阵乘法法则有

$$\begin{pmatrix} \boldsymbol{X}_1 \\ \vdots \\ \boldsymbol{X}_n \end{pmatrix} \boldsymbol{A} = \begin{pmatrix} c_1\boldsymbol{X}_1 \\ \vdots \\ c_n\boldsymbol{X}_n \end{pmatrix} = \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{pmatrix} \begin{pmatrix} \boldsymbol{X}_1 \\ \vdots \\ \boldsymbol{X}_n \end{pmatrix}. \tag{8}$$

这就表明 $\boldsymbol{PA} = \boldsymbol{DP}$，因此 $\boldsymbol{PAP}^{-1} = \boldsymbol{D}$，这里 $\boldsymbol{D}$ 是以 $c_1, \cdots, c_n$ 为对角线元素的对角矩阵. 事实上，矩阵 $\boldsymbol{P}$ 恰好是定理 4 的直接证明中对基变换所需要的矩阵.

证毕

另一方面，存在着与任意对角矩阵都不相似的矩阵 (参看后面的习题 5).

为了明显地构造出与已知矩阵相似的对角矩阵 (如果它存在的话)，我们要找出特征值和特征向量. 根据下面的考虑，特征值和特征向量的求法可以大大简化.

如果标量 λ 是 $n \times n$ 矩阵 $\boldsymbol{A}$ 的特征值，$\boldsymbol{I}$ 是 $n \times n$ 单位矩阵，那么 $\boldsymbol{XA} = \lambda\boldsymbol{X} = \lambda\boldsymbol{XI}$，因此对某个非零 n 维向量 $\boldsymbol{X}$ 有 $\boldsymbol{X}(\boldsymbol{A} - \lambda\boldsymbol{I}) = \boldsymbol{O}$. 于是以 $\boldsymbol{A} - \lambda\boldsymbol{I}$ 为系数矩阵的 n 个齐次线性方程组有非平凡解；因此根据 8.6 节定理 9 的推论 1，我们有

定理 5 标量 λ 是矩阵 $\boldsymbol{A}$ 的特征值当且仅当矩阵 $\boldsymbol{A} - \lambda\boldsymbol{I}$ 是奇异的.

例如，不难看出，2×2 矩阵

$$\boldsymbol{A} - \lambda\boldsymbol{I} = \begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{pmatrix} \tag{9}$$

是奇异的当且仅当

$$\lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21} = 0. \tag{10}$$

(这只表明 $\boldsymbol{A} - \lambda\boldsymbol{I}$ 的行列式等于零.) 因此，我们通过求解这个方程求出所有的特征

值. 而且, 对每个根 λ 至少有一个特征向量, 这可以通过求解方程组

$$x_1a_{11}+x_2a_{21}=\lambda x_1,$$
$$x_1a_{12}+x_2a_{22}=\lambda x_2$$

而得到.

例 求与矩阵 $\begin{pmatrix}-3 & 4\\ 2 & -1\end{pmatrix}$ 相似的对角矩阵.

多项式 (10) 是 $\lambda^2+4\lambda-5$. 这个多项式的根是 1 和 -5; 因此特征向量满足齐次方程组

$$\begin{aligned}-3x+2y&=x,\\ 4x-\ \ y&=y,\end{aligned}\quad \text{或} \quad \begin{aligned}-3x+2y&=-5x,\\ 4x-\ \ y&=-5y.\end{aligned}$$

解这两个方程组, 我们得到特征向量 $(1,2)$ 和 $(1,-1)$. 用这两个向量作新的基, 上述变换的矩阵就呈现对角形. 根据定理 3′, 新的对角矩阵可以写成矩阵的乘积

$$\begin{pmatrix}1 & 2\\ 1 & -1\end{pmatrix}\begin{pmatrix}-3 & 4\\ 2 & -1\end{pmatrix}\begin{pmatrix}1 & 2\\ 1 & -1\end{pmatrix}^{-1}=\begin{pmatrix}1 & 0\\ 0 & -5\end{pmatrix}.$$

习 题

1. 证明: 方程 $2x'=(1+b)x+(1-b)y, 2y'=(1-b)x+(1+b)y$, 表示一个关于通过原点与 x 轴成 45° 角的直线的压缩. 计算这个变换的特征值和特征向量, 并说明它们的几何意义.
2. 计算下列在复数域上的矩阵的特征值和特征向量:

   (a) $\begin{pmatrix}2 & 4\\ 5 & 3\end{pmatrix}$, (b) $\begin{pmatrix}3 & 2\\ -2 & 3\end{pmatrix}$, (c) $\begin{pmatrix}1 & 2\\ 2 & -2\end{pmatrix}$, (d) $\begin{pmatrix}-1 & 2\mathrm{i}\\ -2\mathrm{i} & 2\end{pmatrix}$.
3. 对习题 2 所给出的每个矩阵 $\boldsymbol{A}$, 如果可能的话, 求出非奇异矩阵 $\boldsymbol{P}$, 使得 $\boldsymbol{PAP}^{-1}$ 是对角矩阵.
4. (a) 求出表示转过 θ 角的平面旋转的矩阵的复特征值.

   (b) 证明: 表示转过 θ 角 $(0<\theta<\pi)$ 的平面旋转的矩阵不能同任意实对角矩阵相似.
5. 证明: 当 $c\neq 0$, 矩阵 $\begin{pmatrix}1 & c\\ 0 & 1\end{pmatrix}$ 不能同任意实的或复的对角矩阵相似. 从几何上说明这个结果.
6. 证明: 2×2 矩阵 $\boldsymbol{A}$ 的特征向量的斜率 γ 满足二次方程 $a_{21}\gamma^2+(a_{11}-a_{22})\gamma-a_{12}=0$.
7. 证明: 属于已知矩阵的固定特征值的所有特征向量的集合构成一个子空间, 这时假定 **0** 包含在这些特征向量中.
8. 证明: 非标量矩阵的任意 2×2 实对称矩阵有两个不同的实特征向量.

9. (a) 证明：两个 $m \times n$ 矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是等价的当且仅当它们对于 m 维向量空间 V 与 n 维向量空间 W 的两组不同基表示同一个从 V 到 W 的线性变换 $T: V \to W$.

   (b) 按照这种看法，解释 8.9 节定理 18.

*10. 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都与对角矩阵相似. 证明 $\boldsymbol{AB} = \boldsymbol{BA}$ 当且仅当 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 具有共同的特征向量基 (Frobenius).

*11. (a) 证明：如果 $\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 与一个正交矩阵相似，那么 $ad - bc = \pm 1$. (正交矩阵的定义见 9.4 节.)

   (b) 证明：如果 $ad-bc = 1$, 那么 $\boldsymbol{A}$ 与正交矩阵相似当且仅当 $\boldsymbol{A} = \pm\boldsymbol{I}$ 或 $-2 < a+d < 2$.

   (c) 证明：如果 $ad - bc = -1$, 那么 $\boldsymbol{A}$ 与正交矩阵相似当且仅当 $a + d = 0$.

9.3 全线性群与仿射群

n 维向量空间 F^n 的所有非奇异线性变换构成一个群，因为这样的变换的乘积和逆变换还是线性的和非奇异的 (8.6 节定理 9). 这个群称为*全线性群* $L_n = L_n(F)$. 在线性变换同矩阵的一一对应中，线性变换的乘积对应着矩阵的乘积，所以全线性群与元素属于域 F 的所有 $n \times n$ 非奇异矩阵构成的群同构.

全体平移构成另一个重要的群. 平面上的平移是把平面上的所有点沿着某一指定的方向移动同样一段距离. 移动的距离和方向可以用向量 $\boldsymbol{\kappa}$ 表示，$\boldsymbol{\kappa}$ 具有适当的大小和方向，那么平移把每个向量 $\boldsymbol{\xi}$ 的端点移到向量 $\boldsymbol{\xi} + \boldsymbol{\kappa}$ 的端点. 在空间 F^n 中，平移是变换 $\boldsymbol{\xi} \longmapsto \boldsymbol{\xi} + \boldsymbol{\kappa}$, 其中 $\boldsymbol{\kappa}$ 是固定向量. 对任意坐标系，平移后的向量的坐标是 $y_1 = x_1 + k_1, \cdots, y_n = x_n + k_n$, 这里 k_i 是向量 $\boldsymbol{\kappa}$ 的坐标. 平移 $\boldsymbol{\xi} \longmapsto \boldsymbol{\eta} = \boldsymbol{\xi} + \boldsymbol{\kappa}$ 与 $\boldsymbol{\eta} \longmapsto \boldsymbol{\zeta} = \boldsymbol{\eta} + \boldsymbol{\lambda}$ 的乘积是通过代入而得，它是平移 $\boldsymbol{\xi} \longmapsto \boldsymbol{\zeta} = \boldsymbol{\xi} + (\boldsymbol{\kappa} + \boldsymbol{\lambda})$. 这恰好对应于向量 $\boldsymbol{\kappa}$ 与 $\boldsymbol{\lambda}$ 的和. 类似地，平移 $\boldsymbol{\xi} \longmapsto \boldsymbol{\xi} + \boldsymbol{\kappa}$ 的逆是 $\boldsymbol{\eta} \longmapsto \boldsymbol{\eta} - \boldsymbol{\kappa}$. 于是我们就证明了凯莱定理 (6.5 节定理 8) 的特殊情形：

定理 6 *F^n 的所有平移 $\boldsymbol{\xi} \longmapsto \boldsymbol{\xi} + \boldsymbol{\kappa}$ 构成一个阿贝耳群，这个群与 F^n 的全体向量 $\boldsymbol{\kappa}$ 的加法群同构.*

线性变换 T 后面再跟随一个平移，就得到变换

$$\boldsymbol{\xi} \longmapsto \boldsymbol{\eta} = \boldsymbol{\xi}T + \boldsymbol{\kappa} \quad (T\text{是线性变换, }\boldsymbol{\kappa}\text{是固定向量}). \tag{11}$$

任意一个这种形式的变换称为 F^n 的一个*仿射变换H*. 仿射变换包括线性变换 (当 $\boldsymbol{\kappa} = 0$) 和平移 (当 $T = I$). 如果仿射变换 (11) 后面跟随第二个仿射变换 $\boldsymbol{\eta} \longmapsto \boldsymbol{\eta}U + \boldsymbol{\lambda}$, 则它们的乘积是

$$\boldsymbol{\xi} \longmapsto (\boldsymbol{\xi}T + \boldsymbol{\kappa})U + \boldsymbol{\lambda} = \boldsymbol{\xi}(TU) + (\boldsymbol{\kappa}U + \boldsymbol{\lambda}). \tag{12}$$

某结果还是仿射变换, 因为 $\kappa U+\lambda$ 是 F^n 的一个固定向量. 每个平移是一一的, 也是映上的, 因此它有逆. 所以仿射变换 (11) 是一一映上的当且仅当它的线性部分是一一的. 因此仿射变换 (11) 的逆是仿射变换 $\eta\longmapsto\xi=\eta T^{-1}-\kappa T^{-1}$, 这可通过公式 (11) 解出 ξ 而得到. 这就证明了

定理 7 *F^n 的所有非奇异仿射变换的集合构成一个群, 称为仿射群$A_n(F)$. 全线性群和平移群是它的子群.*

仿射变换对于基的方程是什么? 线性部分 T 产生矩阵 $A=(a_{ij})$; 平移向量按坐标写成行向量 $K=(k_1,\cdots,k_n)$. 于是仿射变换把坐标为 $X=(x_1,\cdots,x_n)$ 的向量变换成坐标为

$$Y=XA+K,\qquad y_j=\sum_{i=1}^{n}x_ia_{ij}+k_j\quad(j=1,\cdots,n)\tag{13}$$

的向量. 一个变换是仿射变换当且仅当它对于某一组基可表示成像 (13) 式那样的非齐次线性方程组.

变换 (13) 与 $Z=YB+L$ 的乘积是

$$Z=X(AB)+KB+L\quad(K,L\text{是行矩阵}).\tag{14}$$

这个公式与 (12) 式相平行. 我们从变换 (13) 按下面方式构造一个 $n+1$ 阶矩阵, 即在矩阵 A 的右边加上一个零列, A 的下面加上行向量 K, 右下方加上单个元素 1:

$$\{Y=XA+K\}\longleftrightarrow\begin{pmatrix}A&O\\K&1\end{pmatrix}(O\text{是 }n\times1\text{ 矩阵, }K\text{ 是 }1\times n\text{ 矩阵}).\tag{15}$$

对于这样的矩阵, 满足同样的乘法法则. 由分块矩阵乘法法则 (8.5 节 (43) 式) 得到

$$\begin{aligned}\begin{pmatrix}A&O\\K&1\end{pmatrix}\begin{pmatrix}B&O\\L&1\end{pmatrix}&=\begin{pmatrix}AB+OL&AO+O\cdot1\\KB+1\cdot L&KO+1\cdot1\end{pmatrix}\\&=\begin{pmatrix}AB&O\\KB+L&1\end{pmatrix}.\end{aligned}\tag{16}$$

等式右边这个结果正是对应于乘积变换 (14) 的加边矩阵. 这就证明了

定理 8 *n 维空间的所有非奇异仿射变换构成的群与所有最后一列为 $(0,\cdots,0,1)^{\mathrm{T}}$ 的 $(n+1)\times(n+1)$ 非奇异矩阵构成的群同构. 这个同构通过对应 (15) 明显地给出.*

每个仿射变换 $\xi H=\xi T+\kappa$ 确定唯一的线性变换 T. 按照 (12) 式, 两个仿射变换的乘积确定相应的线性部分的乘积. 这个对应 $H\longmapsto T$ 把非奇异仿射变换群映上到全线性群, 在群论意义下 (6.11 节), 它是一个同态. 在任意同态中, 映射到单

位元素的元素构成的集合是一正规子群; 这时, 满足 $T=I$ 的仿射变换 H 恰好是一个平移. 这就证明了

定理 9 *平移群是仿射群的正规子群.*

方程 (13) 可以像上述那样解释为点 (向量) 的变换, 它把每个点 $\mathbf{X}=(x_1,\cdots,x_n)$ 变换到同一坐标系中的新点 $\mathbf{Y}=(y_1,\cdots,y_n)$. 同样我们可以把方程 (13) 解释为坐标的变换. 我们称第一种解释为坐标固定图像移动的变换 (即点移动到另一个地方), 称第二种解释为图像固定坐标移动的变换 (即点取另一个名字).

例如, 在平面上, 方程组

$$y_1=x_1+2,\ y_2=x_2-1,$$

当把它作为第一种变换解释时, 它把整个平面向东平移两个单位, 再向南平移一个单位; 当把它作为第二种变换解释时, 原来坐标网格用一个平行的网格来代替, 新的坐标原点是把原来的坐标原点向西移动两个单位再向北移动一个单位而得到.

对所有变换群都可作类似的双重解释.

习　题

1. (a) 用矩阵表示下列各仿射变换:

$$\begin{aligned} H_1: x'&=3x+6y+2, \quad & y'&=3y-4;\\ H_2: x'&=x+y+3, \quad & y'&=x-y+5 \end{aligned}$$

   (b) 计算乘积 H_1H_2, H_2H_1.

   (c) 求 H_1 和 H_2 的逆.

2. 证明: 满足条件 $ad-bc=1$ 的所有仿射变换 $x'=ax+by+e, y'=cx+dy+f$ 的集合是仿射群 $A_2(F)$ 的一个正规子群.

*3. 已知单位圆 $x^2+y^2=1$, 证明: 平面上每个非奇异仿射变换把这个单位圆变成椭圆或圆.

4. 在下列域 F 上 $n\times n$ 矩阵的集合中, 哪些是全线性群的子群?

   (a) 全体标量矩阵 $c\boldsymbol{I}$.　(b) 全体对角矩阵.

   (c) 全体非奇异对角矩阵.　(d) 全体置换矩阵.

   (e) 全体单项矩阵.　(f) 全体三角形矩阵.

   (g) 全体严格三角形矩阵.　(h) 第二行元素为零的全体矩阵.

   (i) 至少有一行的元素为零的全体矩阵.

5. 列举与 F^n 的所有平移构成的群同构的矩阵群.

6. (a) 设 $\mathbf{Z}_2$ 是模 2 整数域, 列出 $L_2(\boldsymbol{Z}_2)$ 中所有矩阵.

   (b) 构造群 $L_2(\mathbf{Z}_2)$ 的乘法表.

*7. 当 $\mathbf{Z}_p$ 是模 p 整数域时, 全线性群 $L_2(\mathbf{Z}_p)$ 的阶是多少?

8. 设 G 是所有形为 $\boldsymbol{A}=\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$(其中 $ad \neq 0$) 的矩阵构成的群. 证明：对应 $\boldsymbol{A} \longmapsto a$ 是一个同态.
9. 将 3×3 非奇异三角形矩阵构成的群同态地映射到 2×2 非奇异三角形矩阵构成的群. (提示：像习题 8 那样证明, 但这里用分块矩阵.)
10. 证明：如果两个域 F 和 K 是同构的, 那么群 $L_n(F)$ 和 $L_n(K)$ 也是同构的.
11. 设 $n<m$, 证明：$L_n(F)$ 与 $L_m(F)$ 的子群同构.
12. (a) 证明：线性群 $L_n(F)$ 的中心由标量矩阵 $c\boldsymbol{I}(c\neq 0)$ 组成. (提示：它们一定同每个矩阵 $\boldsymbol{I}+\boldsymbol{E}_{ij}$ 可交换.)
    (b) 证明：同每个仿射变换可交换的唯一的仿射变换是恒等变换.

*13. 设 $L_n(F)$ 是全线性群, 证明：两个仿射变换 H_1 和 H_2 落入 $L_n(F)$ 的同一个右陪集中当且仅当 $OH_1=OH_2$(O 是原点!).

14. 证明：商群 $A_n(F)/T_n(F)$ 与 $L_n(F)$ 同构, 其中 A_n 表示仿射群, T_n 表示平移群.
15. (a) 证明：满足 $ad\neq bc$ 的全体一一变换 $y=\dfrac{ax+b}{cx+d}$ 构成一个群 (称为线性分式群).
    (b) 证明：这个群与全线性群在模非零标量矩阵的子群之下的商群同构.
    *(c) 对于大于 2×2 的矩阵, 推广上述结果.
16. (a) 证明：所有形为 $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B} \end{pmatrix}$(其中 $\boldsymbol{A}$ 为 $r\times r$ 矩阵, $\boldsymbol{B}$ 为 $s\times s$ 矩阵) 的非奇异矩阵构成的集合与直积 $L_r(F)\times L_s(F)$ 同构.
    (b) 当 $r=2, s=1$, 用上述矩阵确定的 $\mathbf{R}^3$ 的线性变换的几何特征是什么?

9.4 正交群与欧几里得群

在欧几里得几何中, 长度这个概念起着极其重要的作用. 因此我们在欧几里得向量空间中寻求使得所有向量 $\boldsymbol{\xi}$ 的长度 $|\boldsymbol{\xi}|$ 保持不变的那些线性变换.

定义 *欧几里得向量空间的线性变换 T, 如果它保持每个向量 $\boldsymbol{\xi}$ 的长度不变, 即 $|\boldsymbol{\xi}T|=|\boldsymbol{\xi}|$, 那么称 T 为正交变换.*

我们现在来确定欧几里得平面的所有正交变换 $\boldsymbol{Y}=\boldsymbol{X}A$. 因为 $\boldsymbol{A}$ 是正交变换, 所以单位向量 (1,0) 和 (0,1) 的变换式

$$(1,0)\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}=(a_1,a_2),\quad (0,1)\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}=(b_1,b_2) \tag{17}$$

的长度为 1. 根据毕达哥拉斯长度公式, 这就意味着

$$a_1^2+a_2^2=1,\qquad b_1^2+b_2^2=1. \tag{18}$$

此外, 向量 (1,1) 具有长度为 $\sqrt{2}$ 的变换式 (a_1+b_1, a_2+b_2), 所以 $(a_1+b_1)^2+(a_2+$

$b_2)^2=2$, 展开后再把 (18) 代入, 我们得到

$$a_1b_1+a_2b_2=0. \tag{18'}$$

根据 (18), 存在一个角 θ 使得 $\cos\theta=a_1$, $\sin\theta=a_2$. 然后由 (18′) 有 $\tan\theta=\dfrac{a_2}{a_1}=-\dfrac{b_1}{b_2}$, 因此根据 (18) 有 $b_2=\pm\cos\theta, b_1=\mp\sin\theta$. 正负号两种不同的选法恰好给出两个矩阵

$$\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}\quad\begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix} \tag{19}$$

根据 8.1 节公式 (5) 和 (5′), 这两个矩阵分别表示转过 θ 角的旋转和关于与 x 轴成 $\alpha=\dfrac{\theta}{2}$ 角的直线的反射. 因此*每个平面正交变换是旋转或是反射.*

几何上, (19) 式的左边一个正交变换的逆可通过把 θ 用 $-\theta$ 代替而得到; 因此这个逆是原来矩阵的转置. 这个事实 (不像三角公式) 可以推广到 $n\times n$ 正交矩阵.

定理 10 *正交变换 T 具有性质：对每对向量 $\boldsymbol{\xi},\boldsymbol{\eta}$, 有*

(i) *T 保持距离不变, 即 $|\boldsymbol{\xi}-\boldsymbol{\eta}|=|\boldsymbol{\xi}T-\boldsymbol{\eta}T|$.*

(ii) *T 保持内积不变, 即 $(\boldsymbol{\xi},\boldsymbol{\eta})=(\boldsymbol{\xi}T,\boldsymbol{\eta}T)$.*

(iii) *T 保持正交性, 即由 $\boldsymbol{\xi}\perp\boldsymbol{\eta}$ 可推出 $\boldsymbol{\xi}T\perp\boldsymbol{\eta}T$.*

(iv) *T 保持角的大小不变, 即 $\cos\angle(\boldsymbol{\xi},\boldsymbol{\eta})=\cos\angle(\boldsymbol{\xi}T,\boldsymbol{\eta}T)$.*

证明 由于 T 是线性的, 由定义得到 (i). 因为 $\boldsymbol{\xi}\perp\boldsymbol{\eta}$ 意味着 $(\boldsymbol{\xi},\boldsymbol{\eta})=0$, 而且角是通过内积来定义的 (7.9 节 (41)), 所以性质 (iii) 和 (iv) 从性质 (ii) 直接推出. 至于性质 (ii), 从内积的 “双线性” 可证明 $(\boldsymbol{\xi}+\boldsymbol{\eta},\boldsymbol{\xi}+\boldsymbol{\eta})=(\boldsymbol{\xi},\boldsymbol{\xi})+2(\boldsymbol{\xi},\boldsymbol{\eta})+(\boldsymbol{\eta},\boldsymbol{\eta})$. 由这个方程可以利用 “长度”(比如 $|\boldsymbol{\xi}|=(\boldsymbol{\xi},\boldsymbol{\xi})^{\frac{1}{2}}$) 解出 $(\boldsymbol{\xi},\boldsymbol{\eta})$, 形为

$$2(\boldsymbol{\xi},\boldsymbol{\eta})=|\boldsymbol{\xi}+\boldsymbol{\eta}|^2-|\boldsymbol{\xi}|^2-|\boldsymbol{\eta}|^2. \tag{20}$$

由于正交变换 T 保持等式右边的长度不变, 所以它也保持等式左边的内积不变. 这就证明了 (ii). 证毕

反过来, 如果已知变换 T 保持所有的内积不变, 因为长度是通过内积来定义的, 所以 T 一定保持长度不变, 因此它是正交变换.

下面我们要问, 什么样的矩阵对应于正交线性变换? 这个问题至少对于标准正交基的情形是容易回答的.

定理 11 *对于任意标准正交基, $n\times n$ 实矩阵 $\boldsymbol{A}$ 表示一个正交线性变换当且仅当 $\boldsymbol{A}$ 的每个行向量的长度为 1, 任意两个行向量是正交的.*

证明 根据定理 10, 任意正交变换 T 一定把已知基 $\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n$ 变换为一组标准正交基 $\boldsymbol{\alpha}_1=\boldsymbol{\varepsilon}_1T,\cdots,\boldsymbol{\alpha}_n=\boldsymbol{\varepsilon}_nT$. 反过来, 如果变换 T 具有这个性质, 那么对任意向量

$\boldsymbol{\xi} = x_1\boldsymbol{\varepsilon}_1 + \cdots + x_n\boldsymbol{\varepsilon}_n$, 有变换式 $\boldsymbol{\xi}T = x_1\boldsymbol{\alpha}_1 + \cdots + x_n\boldsymbol{\alpha}_n$, 由 7.11 节的定理 22 我们知道, 长度是按普通公式给出

$$|\boldsymbol{\xi}| = (x_1^2 + \cdots + x_n^2)^{\frac{1}{2}} = |\boldsymbol{\xi}T|,$$

因此 T 是正交的. 根据下面的说明 (参看 8.1 节) 我们就完成了定理的证明：即 $\boldsymbol{A}$ 的第 i 行表示向量 $\boldsymbol{\alpha}_i = \boldsymbol{\varepsilon}_i T_{\boldsymbol{A}}$ 关于原基 $\boldsymbol{\varepsilon}_1, \cdots, \boldsymbol{\varepsilon}_n$ 的坐标.

定理中所叙述的关于 $\boldsymbol{A}$ 的条件, 写成坐标形式, 等价于方程

$$\sum_{k=1}^{n} a_{ik}a_{ik} = 1, \quad 对所有\ i, \qquad \sum_{k=1}^{n} a_{ik}a_{jk} = 0, \quad 当\ i \neq j. \tag{21}$$

对于 2×2 矩阵, 结论 (21) 恰好就是在 (18) 和 (18′) 中已经建立的那些公式. 如果我们用 $\boldsymbol{A}_i$ 表示矩阵 $\boldsymbol{A}$ 的第 i 行, 用 $\boldsymbol{A}_i^{\mathrm{T}}$ 表示它的转置, $\boldsymbol{A}_i$ 和 $\boldsymbol{A}_j$ 的内积是矩阵的乘积 $\boldsymbol{A}_i\boldsymbol{A}_j^{\mathrm{T}}$(见 8.5 节的 (34)), 那么条件 (21) 可以写成

$$\boldsymbol{A}_i\boldsymbol{A}_i^{\mathrm{T}} = 1, \quad \boldsymbol{A}_i\boldsymbol{A}_j^{\mathrm{T}} = 0,\ 当\ i \neq j. \tag{21′}$$

在矩阵 $\boldsymbol{A}$ 与它的转置 $\boldsymbol{A}^{\mathrm{T}}$ 的乘积 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ 中, 按照行与列的相乘, (21′) 式表明, 第 i 行乘第 j 列是 $\boldsymbol{A}_i\boldsymbol{A}_j^{\mathrm{T}} = \delta_{ij}$, 这里 δ_{ij} 是单位矩阵 $\boldsymbol{I} = (\delta_{ij})$ 中的第 i 行第 j 列元素, 单位矩阵的对角线元素 $\delta_{ii} = 1$, 其他非对角线元素都是零. (记号 δ_{ij} 称为*克罗内克尔符号*.) 于是我们证明了

定理 12 *一个 $n \times n$ 实矩阵表示正交变换当且仅当 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{I}$.*

方程 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{I}$ 在任意域上都有意义, 因此正交矩阵的概念可以定义得更一般些.

定义 *任意域上的方阵 $\boldsymbol{A}$ 是正交的当且仅当 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{I}$.*

这就意味着正交矩阵 $\boldsymbol{A}$ 的转置 $\boldsymbol{A}^{\mathrm{T}}$ 是 $\boldsymbol{A}$ 的右逆, 因此根据 8.6 节定理 9, 每个正交矩阵 $\boldsymbol{A}$ 是非奇异的, 且满足 $\boldsymbol{A}^{-1} = \boldsymbol{A}^{\mathrm{T}}$. 因此 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{I}$. 这个方程可写成 $\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{I}$, 因此 $\boldsymbol{A}^{\mathrm{T}}$ 是正交矩阵. 这就是说, 任意正交矩阵 $\boldsymbol{A}$ 的转置也是正交矩阵. 由此还可以推出, 矩阵 $\boldsymbol{A}$ 是正交的当且仅当 $\boldsymbol{A}$ 的每一列向量的长度为 1, 任意两个列向量正交：

$$\sum_{k=1}^{n} a_{ki}a_{ki} = 1, \quad 对所有\ i, \qquad \sum_{k=1}^{n} a_{ki}a_{kj} = 0, \quad 当\ i \neq j. \tag{22}$$

所有 $n \times n$ 正交矩阵构成一个群. 这是显然的, 因为正交矩阵的逆 $\boldsymbol{A}^{-1} = \boldsymbol{A}^{\mathrm{T}}$ 是正交的, 并且两个正交矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的乘积是正交的：$(\boldsymbol{A}\boldsymbol{B})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1} = (\boldsymbol{A}\boldsymbol{B})^{-1}$. 这个群是全线性群 $L_n(F)$ 的子群, 称它为*正交群* $O_n(F)$; 当 $F = \mathbf{R}$ 时, $O_n(F)$ 与已知欧几里得空间的所有正交变换构成的群同构.

欧几里得向量空间 E 的刚体运动指的是 E 中保持距离不变的非奇异变换 U, 也就是说, 对所有向量 $\boldsymbol{\xi}, \boldsymbol{\eta}$, U 满足 $|\boldsymbol{\xi} U - \boldsymbol{\eta} U| = |\boldsymbol{\xi} - \boldsymbol{\eta}|$. E 的任意平移保持向量差 $\boldsymbol{\xi} - \boldsymbol{\eta}$ 不变, 因此也保持它们的长度不变, 所以它是刚体运动. 因此, 如果仿射变换 $\boldsymbol{\xi} \longmapsto \boldsymbol{\xi} T + \boldsymbol{\kappa}$ 是刚性变换, 那么 $\boldsymbol{\xi} \longmapsto (\boldsymbol{\eta} - \boldsymbol{\kappa}) = \boldsymbol{\xi} T$ 也是刚性变换. 反过来, 如果 T 是刚性变换, 那么 $\boldsymbol{\xi} \longmapsto \boldsymbol{\eta} = \boldsymbol{\xi} T + \boldsymbol{\kappa}$ 也是刚性变换. 而根据定理 10, 线性变换是刚性的当且仅当它是正交变换, 于是我们得出结论, *仿射变换 (11) 是刚体运动当且仅当 T 是正交变换*. 像在定理 7 的证明中那样, 因为全体正交变换构成群, 所以可以得出, 全体刚性仿射变换构成仿射群的一个子群, 它称为*欧几里得群*. 这是欧几里得几何的基础.①

还有其他各种几何群. 我们所熟悉的一个几何群是所有相似变换 T 构成的群, 它是由使所有长度乘上数因子 $c_T > 0$ 的线性变换 T 组成, 所以 $|\boldsymbol{\xi} T| = c_T|\boldsymbol{\xi}|$. 可以证明, 这些相似变换实际上构成一个群, 它包含正交群作为子群. "广义" 相似群是由所有仿射变换 $\boldsymbol{\xi} \longmapsto \boldsymbol{\xi} T + \boldsymbol{\kappa}$ 组成, 其中 T 是相似变换.

习 题

1. 检验下列矩阵的正交性. 如果某个矩阵是正交的, 求出它的逆矩阵:

$$\text{(a)} \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}, \quad \text{(b)} \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}, \quad \text{(c)} \begin{pmatrix} 0.6 & 0.8 \\ 0.8 & -0.6 \end{pmatrix}.$$

2. 求出一个正交矩阵, 它的第一行是向量 $(5, 12, 0)$ 乘上一个标量.
3. 证明: 如果把正交矩阵的列置换, 那么置换后的矩阵仍是正交矩阵.
4. 证明: 如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是正交矩阵, 那么 $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B} \end{pmatrix}$ 和 $\begin{pmatrix} \boldsymbol{O} & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{O} \end{pmatrix}$ 也是正交矩阵.
5. 把下面两个矩阵相乘, 检验乘积矩阵的正交性:

$$\begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}.$$

6. 证明: 欧几里得群与矩阵群同构.
7. 证明: 全体平移构成欧几里得群的正规子群.
8. 作为定理 10 中的性质 (ii) 的另一个证明, 用基本原理证明 $4(\boldsymbol{\xi}, \boldsymbol{\eta}) = |\boldsymbol{\xi} + \boldsymbol{\eta}|^2 - |\boldsymbol{\xi} - \boldsymbol{\eta}|^2$.
9. 证明: 一个仿射变换 H 同每个平移可交换当且仅当 H 本身是一个平移.
10. 证明: 任意相似变换 S 只能按照一种方法写成形式 $S = cT$, 它是正标量 c 和正交变换 T 的乘积.

① 事实上, 任意刚体运动必须是仿射变换, 因此欧几里得群实际上是所有刚体运动的群.

11. 对于标准正交基, 给出矩阵 $\boldsymbol{A}$ 表示相似变换的充分必要条件. (参看定理 11 和定理 12.)
12. (a) 证明：全体相似变换构成群 S_n.
    (b) 证明：O_n 是 S_n 的正规子群.
    (c) 证明：商群 S_n/O_n 与全体正实数构成的乘法群同构.
13. $\mathbf{Z}_2$ 上的 3×3 矩阵中有多少个正交矩阵? $\mathbf{Z}_3$ 上呢?
14. (a) 证明：对应 $\boldsymbol{A}\longmapsto\theta(\boldsymbol{A})=(\boldsymbol{A}^{-1})^{\mathrm{T}}$ 是全线性群 $L_n(F)$ 的一个自同构.
    (b) 证明：对所有的 $\boldsymbol{A}$, 有 $\theta^2(\boldsymbol{A})=\boldsymbol{A}$.
    (c) 对哪些矩阵 $\boldsymbol{A}$, 有 $\theta(\boldsymbol{A})=\boldsymbol{A}$?

9.5 不变量与标准型

全线性群、仿射群、正交群和欧几里得群都是线性群的例子. 另一个例子是酉群 (9.12 节). 在下面几节中, 我们将看到使用这些群中适当的变换, 多项式、二次型以及各种几何图形可以 "简化" 到什么程度. 这些化简类似于化简一般矩阵为行等价的简化梯形矩阵, 已经证明了它们的秩在所考虑的变换之下是一个不变量. 化简后的 "标准型" 和 "不变量" 的概念可以给出下面更一般的描述.[①]

设 G 是任意集合或 "空间"S 上的变换群 (6.2 节). 我们称 S 上两个元素 x 和 y 在 G 之下是**等价的** (记作 xE_Gy) 当且仅当存在 G 的某一变换 T 把 x 变到 y. 那么 T^{-1} 把 y 变回 x, 所以 yE_Gx, 因此这个等价关系是对称的. 类似地, 用群的其他性质我们可以证明, 在任意 G 之下的等价性还满足自反律和传递律 (即为等价关系). 对于 S 的子集合 C, 如果每个元素 $x\in S$, 在 G 之下与 C 中一个且只有一个元素 c 等价, 那么 C 称为 S 在 G 之下的**标准型集合**; 这个元素 c 就是 x 的**标准型**. 对 S 的所有元素 x 定义的函数 $F(x)$, 它取值在另一个适当的集合上, 比如说一个数集, 如果它对于 S 中每个点 x 和 G 中每个变换 T 有 $F(xT)=F(x)$, 则称函数 $F(x)$ 是 G 之下的**不变量**(有时称为不变式); 换句话说, F 在所有等价元素上的值一定是相同的. 一组不变量 $F_1,\cdots,F_n$, 如果可由 $F_1(x)=F_1(y),\cdots,F_n(x)=F_n(y)$ 推出 x 和 y 等价, 则称它们是 G 之下的**全系不变量**.

例如, 设空间 S 是某一域上所有 $n\times n$ 矩阵组成的集合 M_n. 对这些矩阵我们手边已经有三种不同等价关系, 连同三种新的等价关系一起列在下面. 后面三种新的等价关系将在以后几节 (分别在 9.8 节、9.10 节和 9.12 节) 里讨论.

$\boldsymbol{A}$ 行等价于 $\boldsymbol{B}$	$\boldsymbol{B}=\boldsymbol{PA}$,	$\boldsymbol{P}$ 为非奇异矩阵;
$\boldsymbol{A}$ 等价于 $\boldsymbol{B}$	$\boldsymbol{B}=\boldsymbol{PAQ}$,	$\boldsymbol{P},\boldsymbol{Q}$ 为非奇异矩阵;
$\boldsymbol{A}$ 相似于 $\boldsymbol{B}$	$\boldsymbol{B}=\boldsymbol{PAP}^{-1}$,	$\boldsymbol{P}$ 为非奇异矩阵;

① 读者注意, 当你读完本章之后, 再回来讨论这一般情形.

$\boldsymbol{A}$ 相合于 $\boldsymbol{B}$	$\boldsymbol{B}=\boldsymbol{PAP}^{\mathrm{T}}$,	$\boldsymbol{P}$ 为非奇异矩阵;
$\boldsymbol{A}$ 正交等价于 $\boldsymbol{B}$	$\boldsymbol{B}=\boldsymbol{PAP}^{-1}$,	$\boldsymbol{P}$ 为正交矩阵;
$\boldsymbol{A}$ 酉等价于 $\boldsymbol{B}$	$\boldsymbol{B}=\boldsymbol{PAP}^{-1}$,	$\boldsymbol{P}$ 为酉矩阵.

上面第一个等价关系读作: "$\boldsymbol{A}$ 行等价于 $\boldsymbol{B}$ 当且仅当存在非奇异矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{B}=\boldsymbol{PA}$". 对其他各关系可给出类似的读法.

上述每个等价关系是由作用在 M_n 上的一个适当的群 G 所确定的等价关系 E_G, 而且自然产生于对矩阵的一种解释.

第一个行等价关系是在讨论把 F^n 的一个固定子空间表示成矩阵 $\boldsymbol{A}$ 的行空间中产生的. 在这种情况下, 矩阵 $\boldsymbol{P}$ 的全线性群通过 $\boldsymbol{A}\longmapsto\boldsymbol{PA}$ 作用在 $\boldsymbol{A}$ 上, 简化梯形矩阵是这个群之下的标准型. $\boldsymbol{A}$ 的秩是这个群之下的不变量 (数值的), 但它并没有给出全系不变量, 这因为秩相同的两个矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 不一定是行等价的.

第二个等价关系 (按 $\boldsymbol{B}=\boldsymbol{PAQ}$ 这种意义的等价, 不会同一般的等价关系混淆) 是在讨论一个向量空间到另一个向量空间的线性变换的各种矩阵表示中产生的 (参看 9.2 节习题 9). 这里, 根据 8.9 节的定理 18, 秩是在群 $\boldsymbol{A}\longmapsto\boldsymbol{PAQ}$ 之下的全系不变量. 对角线元素是 1 或 0(1 都在 0 的前面) 的所有对角矩阵组成的集合是标准型集合. 注意, 我们同样还可以选取不同的标准型集合, 比如说它们是同一类型的对角矩阵, 只是对角线上的 0 在 1 的前面.

相似关系是在讨论一个向量空间到自身的线性变换的各种矩阵表示中产生的, 在这种情形, 全线性群是通过 $\boldsymbol{A}\longmapsto\boldsymbol{PAP}^{-1}$ 作用到 $\boldsymbol{A}$ 上. 在相似变换之下, 矩阵 $\boldsymbol{A}$ 的秩是一个不变量, 因为两个相似矩阵是等价的, 在等价之下, 秩是不变量. 根据 9.2 节, 矩阵的所有特征值集合在相似变换之下也是不变量, 但它不是全系不变量. 在相似变换之下, 系统地列举出完全的标准型集合是矩阵论中一个重要的问题; 对于复数域, 给出了矩阵的若当标准型 (见 10.10 节).

后面将出现的相合关系 ($\boldsymbol{B}=\boldsymbol{PAP}^{\mathrm{T}}$), 它是在讨论用 (对称) 矩阵表示二次型中产生的.

作为在群之下等价的另一个例子, 我们考虑用所有平移 $y=x+k$ 构成的群来化简二次多项式 $f(x)=ax^2+bx+c$, 其中 $a\neq 0$. 将 $x=y-k$ 代入, 我们得到平移 $f(x)$ 的结果是

$$g(y)=a(y-k)^2+b(y-k)+c=ay^2+(b-2ak)y+ak^2-bk+c.$$

特别是, 我们得到熟知的 "配方" 法 —— 所得的新多项式没有一次项当且仅当 $k=\dfrac{b}{2a}$, 这时, 多项式是

$$g(y)=ay^2-\frac{d}{4a},\quad \text{其中}\quad d=b^2-4ac. \tag{23}$$

于是 $f(x)$ 在平移群之下与一个且只与一个形为 ay^2+h 的多项式等价, 因此没有一次项的二次多项式是这个群之下的标准型. 另一方面, 任意平移后的多项式同原多项式 $f(x)$ 具有相同的首项系数 a 和相同的判别式 $d=(b-2ak)^2-4a(ak^2-bk+c)$. 因此 $f(x)$ 的首项系数和判别式是这个群之下的不变量, 它们组成全系不变量, 因为标准型可以通过首项系数和判别式表示成 (23).

为了给出最后一个例子, 回忆一下, 全线性群 $L_n(F)$ 是向量空间 F^n 的变换群. 这个群的每个变换把 F^n 的子空间 S 变到另一个子空间. 根据 8.6 节定理 10 的推论 2, 任意子空间 S 的维数是全线性群之下的不变量. 这一个不变量实际上是全线性群之下关于 F^n 的子空间的全系不变量 (见下面习题 5).

习　题

1. 求出所有的首项系数为 1 的二次多项式 x^2+bx+c 在平移群之下的标准型.
2. 求出所有二次多项式 ax^2+bx+c(其中 $a\neq 0$) 在仿射群 $y=hx+k, h\neq 0$ 之下的标准型.
3. 在习题 2 中, 证明 $\dfrac{d}{a}=\dfrac{b^2}{a}-4c$ 是一个仿射不变量.
4. 证明: 在满足条件 $1+1\neq 0$ 的任意域上, 任意四次多项式在平移之下等价于一个没有三次项的多项式.
5. 设 V 是 n 维向量空间, 证明: V 的子空间的有序对 (S_1,S_2) 在全线性群之下的全系不变量是由 S_1,S_2 的维数以及 S_1,S_2 的交的维数给出.
6. 考虑在群 $x\longmapsto rx$(其中 $r\neq 0$ 为有理数) 之下的含有有理系数 a 的齐次二次函数 ax^2 的集合. 证明: 系数 a 为不同素数乘积 (无平方因子) 的二次函数的集合提供了上述二次函数集合的标准型.
7. 设 $f(x)$ 是一个变量的任意多项式, 证明: $f(x)$ 的次数和实根的个数是在仿射群之下的两个不变量.
8. 证明: 对于 n 个变量的多项式, 最高次项的系数是平移群之下的不变量.
9. 证明: 一个实三次多项式在仿射群之下与一个且只与一个形为 x^3+ax+b 的多项式等价.

*10. 考虑二次函数 x^2+bx+c, 其中系数 b,c 是模 2 整数域 $\mathbf{Z}_2$ 中的元素, 求出这种二次函数在平移群之下的标准型.

9.6　线性型与双线性型

域 F 上的 n 个变量的线性型是形为

$$f(x_1,\cdots,x_n)=b_1x_1+\cdots+b_nx_n+c \tag{24}$$

的多项式, 其中系数 $b_1,\cdots,b_n$ 和 c 都在 F 中. 除了平凡情形之外, 我们可假定某一

个系数 b_j 不为零. 如果 $c=0$, 则这个线性型称为*齐次线性型*. 任意线性型 (24) 可以看作 F^n 的向量 $\boldsymbol{X}=(x_1,\cdots,x_n)$ 的函数 $f(\boldsymbol{X})$. 不同的线性型确定不同的函数, 这因为函数 $f(\boldsymbol{X})$ 通过公式

$$f(0,\cdots,0)=c,\ f(1,0,\cdots,0)=b_1+c,\ \cdots,\ f(0,\cdots,0,1)=b_n+c$$

确定线性型 (24) 的系数.

对任意线性型, 我们可以应用非奇异仿射变换

$$x_i=\sum_j a_{ij}y_j+k_i,\ (a_{ij})\ \text{是非奇异的} \tag{25}$$

把它代入 (24) 中, 产生新的线性型

$$g(y_1,\cdots,y_n)=\sum_j\Big(\sum_i b_ia_{ij}\Big)y_j+\Big(\sum_i b_ik_i+c\Big). \tag{26}$$

如果存在一个这样的仿射变换把 f 变到 g, 我们就说 f 和 g 在仿射群之下是*等价线性型*.

不难得到线性型的标准型. 首先, 因为某个 $b_j\neq 0$, 平移 $x_j=y_j-\dfrac{c}{b_j}, x_i=y_i$(当 $i\neq j$) 将消去常数项. 置换 $z_1=y_j, z_j=y_1, z_i=y_i$(当 $i\neq 1$ 和 $i\neq j$) 将给出像 (24) 那样一个新的线性型, 其中 $b_1\neq 0$, 且 $c=0$. 如果这个线性型按变量 $x_1,\cdots,x_n$ 写出, 那么由方程

$$y_1=b_1x_1+\cdots+b_nx_n,\quad y_2=x_2,\ \cdots,\ y_n=x_n$$

给出的新的仿射变换是非奇异的, 它把满足 $c=0$ 的任意函数 f 变到等价的函数 $g(y_1,\cdots,y_n)=y_1$. 因此, *在仿射群之下, 所有非零线性型是等价的*.

现在考虑在欧几里得群 (即 (25) 式中的 $\boldsymbol{A}=(a_{ij})$ 是正交矩阵) 之下实线性型的等价性. $d=(b_1^2+\cdots+b_n^2)^{\frac{1}{2}}$ 称为线性型 (24)的*范数*. 如上所述, 通过平移我们可以消去常数 c. 适当选取 d 使得 $\left(\dfrac{b_1}{d},\cdots,\dfrac{b_n}{d}\right)$是具有单位长度的向量, 因此存在一个正交矩阵 (h_{ij}), 以上述向量作为它的第一行. 那么变换 $y_i=\Sigma h_{ij}x_j$ 属于欧几里得群; 因为 $dy_1=b_1x_1+\cdots+b_nx_n$, 所以它把满足 $c=0$ 的线性型 f 变到线性型 $g=dy_1$.

这个线性型 dy_1 是线性型在欧几里得群之下的标准型. 为了证明这一点, 我们只须证明范数 d 是在欧几里得群下的不变量. f 的范数 d 正好是系数向量 $\boldsymbol{\beta}=(b_1,\cdots,b_n)$ 的长度, (26) 式表明, 变换后的线性型中的系数向量是原来的系数向量在正交矩阵 (a_{ij}) 作用之下的变换式 $\boldsymbol{\beta A}$; 因此范数是不变量. 于是我们证明了

定理 13 在欧几里得群之下, 每个线性型 (24) 与一个且只与一个标准型 dy 等价, 其中 d 是满足 $d=(b_1^2+\cdots+b_n^2)^{\frac{1}{2}}$ 的正数, 它是这个群之下的不变量.

关于两组变量 $x_1,\cdots,x_m$ 和 $y_1,\cdots,y_n$ 的 (齐次)双线性型是形为

$$b(x_1,\cdots,x_m,y_1,\cdots,y_n)=\sum_{i=1}^{m}\sum_{j=1}^{n}x_i a_{ij} y_j \tag{27}$$

的多项式, 它是通过系数矩阵 $\boldsymbol{A}=(a_{ij})$ 来确定的. 这个双线性型可以利用向量 $\boldsymbol{X}=(x_1,\cdots,x_m)$ 和 $\boldsymbol{Y}=(y_1,\cdots,y_n)$ 写成矩阵乘积

$$b(\boldsymbol{X},\boldsymbol{Y})=\boldsymbol{X}\boldsymbol{A}\boldsymbol{Y}^{\mathrm{T}}. \tag{28}$$

作为 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 的函数, 这个函数分别对每个自变量是线性的.

更一般地, 设 V 和 W 是同一个域 F 上的维数分别为 m 和 n 的有限维向量空间, 且设 $B(\boldsymbol{\xi},\boldsymbol{\eta})$ 是对于自变量 $\boldsymbol{\xi}\in V$ 和 $\boldsymbol{\eta}\in W$ 定义的取值在 F 上的任意函数, 它按下述意义是双线性的: 对于 a_1 和 $a_2\in F$, 有

$$B(a_1\boldsymbol{\xi}_1+a_2\boldsymbol{\xi}_2,\boldsymbol{\eta})=a_1B(\boldsymbol{\xi}_1,\boldsymbol{\eta})+a_2B(\boldsymbol{\xi}_2,\boldsymbol{\eta}),\quad \boldsymbol{\xi}_1,\boldsymbol{\xi}_2\in V,\quad \boldsymbol{\eta}\in W; \tag{29}$$

$$B(\boldsymbol{\xi},a_1\boldsymbol{\eta}_1+a_2\boldsymbol{\eta}_2)=B(\boldsymbol{\xi},\boldsymbol{\eta}_1)a_1+B(\boldsymbol{\xi},\boldsymbol{\eta}_2)a_2,\quad \boldsymbol{\xi}\in V,\quad \boldsymbol{\eta}_1,\boldsymbol{\eta}_2\in W. \tag{29$'$}$$

选取 V 的基 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 和 W 的基 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n$, 且设 a_{ij} 是按 $a_{ij}=B(\alpha_i,\beta_j)$ 定义的标量. 那么对 V 和 W 中任意向量 $\boldsymbol{\xi}$ 和 $\boldsymbol{\eta}$, 按照基表示, 我们有

$$B(\boldsymbol{\xi},\boldsymbol{\eta})=B(x_1\alpha_1+\cdots+x_m\alpha_m,y_1\beta_1+\cdots+y_n\beta_n).$$

因此根据 (29) 和 (29$'$) 得

$$B(\boldsymbol{\xi},\boldsymbol{\eta})=\sum_{i,j}x_iB(\alpha_i,\beta_j)y_j=\sum_{i,j}x_ia_{ij}y_j.$$

换句话说, V 和 W 上的任意双线性函数 B 对于给定的一组基有像 (27) 那样的唯一表达式. 按照 8.5 节的记号, 另一等价的说法是, 双线性型恰好是 m 维行向量 $\boldsymbol{X}$, $m\times n$ 矩阵 $\boldsymbol{B}$ 和 n 维列向量 $\boldsymbol{Y}$ 的乘积 $\boldsymbol{XBY}$.

这两个空间的基变换对应着各组变量的非奇异变换 $\boldsymbol{X}=\boldsymbol{X}^*\boldsymbol{P}$ 和 $\boldsymbol{Y}=\boldsymbol{Y}^*\boldsymbol{Q}$. 在这些变换下, 可用新的双线性型 $\boldsymbol{X}^*(\boldsymbol{PAQ}^{\mathrm{T}})\boldsymbol{Y}^{*\mathrm{T}}$ 代替 (28) 式, 其中 $\boldsymbol{PAQ}^{\mathrm{T}}$ 是一个新的矩阵. 因为任意非奇异矩阵可以写成一个非奇异矩阵的转置 $\boldsymbol{Q}^{\mathrm{T}}$, 所以我们看出, 两个双线性型 (在基变换之下) 是等价的当且仅当它们的矩阵是等价的. 因此, 根据 8.9 节中关于矩阵等价性的定理 18, 可知任意双线性型与一个且只与一个标准型

$$x_1y_1+\cdots+x_ry_r$$

等价. 这里整数 r 是双线性型的矩阵的秩, 它是一个 (全系) 不变量.

习 题

1. 求出齐次实线性函数在相似群之下的标准型.
2. 分别在下面两种情形下求齐次实线性函数的标准型:
   (a) 在对角线变换群之下, 即 $y_1 = d_1x_1, \cdots, y_n = d_nx_n$;
   (b) 在单项变换群之下, 即 $\boldsymbol{Y} = \boldsymbol{XM}$, 其中 $\boldsymbol{M}$ 为单项矩阵.
3. 证明: 秩为 r 的任意双线性型可以表示成
$$\sum_{i=1}^{r}(b_{i1}x_1 + \cdots + b_{in}x_n)(c_{i1}y_1 + \cdots + c_{in}y_n),$$
   即表示成 r 个线性型的乘积之和.
4. 求出两组新变量 x^*, y^*, z^* 和 u^*, v^*, w^*, 把下面双线性函数化为标准型:
$$xu + xv + xw + yu + yv + yw + zu + zv + zw.$$

9.7 二 次 型

下面 4 节专门研究二次型在各种变换群之下的标准型. 这类问题中最简单的是产生于对平面有心二次曲线(具有 "斜" 轴的椭圆或双曲线) 的研究. 这样的二次曲线满足方程 $\boldsymbol{A}x^2 + \boldsymbol{B}xy + \boldsymbol{C}y^2 = 1$, 其中左边是 "二次型". 这样的二次型 (全体变量的二次齐式) 产生于很多其他情形: 例如, 空间的二次曲面方程, 二次曲线在齐次坐标下的射影方程, 向量长度的平方公式 $|\boldsymbol{X}|^2 = x_1^2 + x_2^2 + \cdots + x_n^2$, 具有三个速度分量 u, v, w 的空间运动物体的动能公式 $\dfrac{m}{2}(u^2 + v^2 + w^2)$, 微分几何中在球面坐标下空间的弧长 ds 的公式 $ds^2 = dr^2 + r^2d\phi^2 + r^2\sin\phi d\theta^2$.

这样的二次型可以用矩阵表示. 例如, 为了得到二次型 $5x^2 + 6xy + 2y^2$ 的矩阵表示, 首先调整二次型使得 xy 和 yx 的系数相等, 写成 $5x^2 + 3xy + 3yx + 2y^2$. 这个表达式可以写成矩阵乘积
$$(x, y)\begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = (x, y)\begin{pmatrix} 5x + 3y \\ 3x + 2y \end{pmatrix} = 5x^2 + 6xy + 2y^2.$$
这里给出的 2×2 系数矩阵是对称的. 由于对称性, 这个矩阵等于它的转置.

一般地, 如果方阵 $\boldsymbol{A}$ 等于它的转置: $\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}$, 则称 $\boldsymbol{A}$ 是对称的; 换句话说, (a_{ij}) 是对称的当且仅当对所有的 i, j, 有 $a_{ij} = a_{ji}$. 类似地, 如果矩阵 $\boldsymbol{C}$ 满足 $\boldsymbol{C}^{\mathrm{T}} = -\boldsymbol{C}$, 则称 $\boldsymbol{C}$ 是斜对称的. 为把矩阵 $\boldsymbol{B}$ 分成对称部分和斜对称部分, 我们可以把 $\boldsymbol{B}$ 写成
$$\boldsymbol{B} = \frac{\boldsymbol{B} + \boldsymbol{B}^{\mathrm{T}}}{2} + \frac{\boldsymbol{B} - \boldsymbol{B}^{\mathrm{T}}}{2} = \boldsymbol{S} + \boldsymbol{K}, \tag{30}$$

其中 $S=\dfrac{B+B^{\mathrm{T}}}{2},K=\dfrac{B-B^{\mathrm{T}}}{2}$. 根据转置的运算法则, 有 $(B\pm B^{\mathrm{T}})^{\mathrm{T}}=B^{\mathrm{T}}\pm(B^{\mathrm{T}})^{\mathrm{T}}=B^{\mathrm{T}}\pm B$, 所以 S 是对称的, K 是斜对称的. 不可能再有别的分解使得 $B=S_1+K_1$, 其中 S_1 是对称的, K_1 是斜对称的. 这是因为任何这样的分解将给出 $B^{\mathrm{T}}=S_1^{\mathrm{T}}+K_1^{\mathrm{T}}=S_1-K_1$, 于是 $B+B^{\mathrm{T}}=2S_1, B-B^{\mathrm{T}}=2K_1$, 所以 $S_1=S, K_1=K$. 公式 (30) 可以应用到任意满足条件 $2=1+1\neq 0$ 的域上, 但对于域 $\mathbf{Z}_2$ 上的矩阵, 公式 (30) 没有意义, 因为在 $\mathbf{Z}_2$ 中 $1+1=0$. 总之, 任意矩阵可以唯一地表示成对称矩阵与斜对称矩阵之和, 只要假定 $1+1\neq 0$.

n 个变量 $x_1,\cdots,x_n$ 的*齐次二次型*是由多项式

$$\sum_i\sum_j x_i b_{ij} x_j$$

来定义的, 其中每一项的次数都是 2. 这个二次型可以写成矩阵的乘积 XBX^{T}. 如果系数矩阵 B 是斜对称的, 那么 $b_{ij}=-b_{ji}$, 因此这个二次型等于零. 一般地, 根据 (30) 式, 把 B 写成 $B=S+K$, 二次型变成

$$XBX^{\mathrm{T}}=X(S+K)X^{\mathrm{T}}=XSX^{\mathrm{T}}+XKX^{\mathrm{T}}=XSX^{\mathrm{T}},\ (K\text{ 是斜对称矩阵}).$$

因此, 若 $1+1\neq 0$, 则任意二次型可以唯一地表示成 (用 A 记 S)

$$\sum_{i=1}^{n}\sum_{j=1}^{n} x_i a_{ij} x_j = XAX^{\mathrm{T}},\quad A=(a_{ij})\text{ 是对称矩阵}. \tag{31}$$

如果向量 ξ 有坐标 $X=(x_1,\cdots,x_n)$, 那么每个二次型确定一个向量 ξ 的二次函数 $Q(\xi)=XAX^{\mathrm{T}}$. 由空间的基变换给出新坐标 X^*, 它与老坐标的关系是通过方程 $X=X^*P$(其中 P 是非奇异矩阵) 给出的. 按照 ξ 的新坐标来表示, 二次型变成

$$Q(\xi)=XAX^{\mathrm{T}}=(X^*P)A(X^*P)^{\mathrm{T}}=X^*(PAP^{\mathrm{T}})X^{*\mathrm{T}};$$

这是含有新矩阵 PAP^{T} 的另一个二次型. 这个新矩阵同 A 一样是对称的, 因为 $(PAP^{\mathrm{T}})^{\mathrm{T}}=(P^{\mathrm{T}})^{\mathrm{T}}A^{\mathrm{T}}P^{\mathrm{T}}=PAP^{\mathrm{T}}$.

定理 14 *通过坐标变换, 含有矩阵 A 的二次型变为含有矩阵 PAP^{T} 的二次型, 这里 P 是非奇异的.*

对称矩阵 A 和 B, 如果满足关系 $B=PAP^{\mathrm{T}}$(其中 P 是非奇异的), 则称 A 和 B *是相合的*.

再重复一下, 定理 14 断言, 齐次的二次型在全线性群 (这个群由关于变量的非奇异的线性齐次变换构成) 之下化为标准型的问题, 等价于求对称矩阵 A 在变换群 $A\longmapsto PAP^{\mathrm{T}}$ 之下的标准型问题.

习 题

1. 证明：$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 和 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ 总是对称的.
2. 证明：如果 $\boldsymbol{A}$ 是斜对称的, 那么 $\boldsymbol{A}^2$ 是对称的.
3. 把 8.3 节中习题 1 的每个矩阵表示成 $\boldsymbol{S}+\boldsymbol{K}$ 的形式.
4. 求出与下列各二次型相联系的对称矩阵：

   (a) $2x^2+3xy+6y^2$, (b) $8xy+4y^2$,

   (c) $x^2+2xy+4xz+3y^2+yz+7z^2$, (d) $4xy$,

   (e) $x^2+4xy+4y^2+2xz+z^2+4yz$.
5. (a) 证明：如果 $\boldsymbol{S}$ 是对称的, $\boldsymbol{A}$ 是正交的, 那么 $\boldsymbol{A}^{-1}\boldsymbol{S}\boldsymbol{A}$ 是对称的.

   (b) 证明：如果 $\boldsymbol{K}$ 是斜对称的, $\boldsymbol{A}$ 是正交的, 那么 $\boldsymbol{A}^{-1}\boldsymbol{K}\boldsymbol{A}$ 是斜对称的.
6. 讨论矩阵 $\boldsymbol{AB}-\boldsymbol{BA}$ 在下列各情况下的对称性：

   (a) $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是对称的. (b) $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是斜对称的.

   (c) $\boldsymbol{A}$ 是对称的, 而 $\boldsymbol{B}$ 是斜对称的.
7. 证明：如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是对称的, 那么 $\boldsymbol{AB}$ 是对称的当且仅当 $\boldsymbol{AB}=\boldsymbol{BA}$.
8. (a) 证明：域 $\mathbf{Z}_2$(模 2 整数域) 上每个斜对称矩阵是对称的.

   (b) 举出 $\mathbf{Z}_2$ 上的矩阵, 它不能表示成和 $\boldsymbol{S}+\boldsymbol{K}$(见 (30) 式).
9. 设 $\boldsymbol{D}$ 是无重复元素的对角矩阵, 证明：$\boldsymbol{AD}=\boldsymbol{DA}$ 当且仅当 $\boldsymbol{A}$ 也是对角矩阵.
10. 设 $Q(\boldsymbol{\xi})$ 是二次函数, 证明

$$Q(\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{\gamma})-Q(\boldsymbol{\alpha}+\boldsymbol{\beta})-Q(\boldsymbol{\beta}+\boldsymbol{\gamma})-Q(\boldsymbol{\gamma}+\boldsymbol{\alpha})+Q(\boldsymbol{\alpha})+Q(\boldsymbol{\beta})+Q(\boldsymbol{\gamma})=0.$$

11. 双线性型 $B(\boldsymbol{\xi},\boldsymbol{\eta})(\boldsymbol{\xi}\in V\boldsymbol{\eta}\in V)$ 如果满足 $B(\boldsymbol{\xi},\boldsymbol{\eta})=B(\boldsymbol{\eta},\boldsymbol{\xi})$, 则称它是对称的. 证明：如果 B 是对称双线性型, 那么 $Q(\boldsymbol{\xi})=B(\boldsymbol{\xi},\boldsymbol{\xi})$ 是满足关系式

$$2B(\boldsymbol{\xi},\boldsymbol{\eta})=Q(\boldsymbol{\xi}+\boldsymbol{\eta})-Q(\boldsymbol{\xi})-Q(\boldsymbol{\eta})$$

的二次型.

12. 证明：$n\times n$ 实矩阵 $\boldsymbol{A}$ 是对称的当且仅当与它相联系的 n 维欧几里得空间的线性变换 $T=T_{\boldsymbol{A}}$ 对任意两个向量 $\boldsymbol{\xi},\boldsymbol{\eta}$ 满足关系 $(\boldsymbol{\xi}T,\boldsymbol{\eta})=(\boldsymbol{\xi},\boldsymbol{\eta}T)$.

*13. 证明：如果实矩阵 $\boldsymbol{S}$ 是斜对称的, 并且 $\boldsymbol{I}+\boldsymbol{S}$ 是非奇异的, 那么 $(\boldsymbol{I}-\boldsymbol{S})(\boldsymbol{I}+\boldsymbol{S})^{-1}$ 是正交矩阵.

9.8 全线性群之下的二次型

大家熟悉的“配方”法可以用作为化简二次型 (通过线性变换) 的方法. 对于两个变量的二次型, 由这个方法可以得到

$$ax^2+2bxy+cy^2$$

$$= a\left(x^2 + 2\frac{b}{a}xy + \frac{b^2}{a^2}y^2\right) + \left(c - \frac{b^2}{a}\right)y^2$$
$$= a\left(x + \frac{b}{a}y\right)^2 + \left(c - \frac{b^2}{a}\right)y^2.$$

括号中的项给出新的变量 $x' = x + \frac{b}{a}y, y' = y$. 在这个变量线性变换下, 二次型变成 $ax'^2 + \left(c - \frac{b^2}{a}\right)y'^2$, 交叉项被去掉了.

这一论证要求 $a \neq 0$. 如果 $a = 0$, 而 $c \neq 0$, 则可用类似的变换来化简. 最后, 如果 $a = c = 0$, 那么原来的二次型是 $2bxy$, 对应的方程 $2bxy = 1$ 表示一个等轴双曲线. 在这种情况下, 变换 $x = x' + y', y = x' - y'$ 将把二次型化为

$$2b(x' + y')(x' - y') = 2b(x'^2 - y'^2).$$

这个表达式也只包含着平方项. (提示: 这里使用的变换与关于双曲线轴的旋转有什么关系?)

类似的 "配方" 法可以应用到多于两个变量的二次型上.

引理 通过非奇异线性变换, 任意不恒为零的二次型 $\Sigma x_i a_{ij} x_j$ 可以化为首项系数 $a_{11} \neq 0$ 的二次型, 只要假定 $1 + 1 \neq 0$.

证明 根据假设, 至少有一个系数 $a_{ij} \neq 0$. 如果是对角线元素 $a_{ii} \neq 0$, 那么我们通过 x_1 与 x_i 变量对换 (这是一个非奇异变换, 因为它的矩阵是置换矩阵), 可以得到新的系数 $a'_{11} \neq 0$. 如果所有对角线元素 a_{ii} 都是零, 但是存在下标 $i \neq j$ 满足 $a_{ij} \neq 0$. 通过置换变量, 我们可以使 $a_{12} \neq 0$; 由矩阵的对称性有 $a_{12} = a_{21}$. 那么已知的二次型就是 $a_{12}x_1x_2 + a_{21}x_2x_1 = 2a_{12}x_1x_2$ 再加上含有其他变量的一些项. 恰好同等轴双曲线的情形一样, 通过变换

$$x_1 = y_1 - y_2, \quad x_2 = y_1 + y_2, \quad x_3 = y_3, \quad \cdots, \quad x_n = y_n,$$

这个二次型就可化成首项系数 $2a_{12} \neq 0$ 的形式 $2a_{12}(y_1^2 - y_2^2)$. 上述变换是非奇异的, 由消去法我们容易证明它有逆变换

$$y_1 = \frac{x_1 + x_2}{2}, \quad y_2 = \frac{x_2 - x_1}{2}, \quad y_3 = x_3, \quad \cdots, \quad y_n = x_n.$$

问: 这个论证中什么地方用到假设 $1 + 1 \neq 0$?

现在我们对任意二次型 "配方". 根据引理, 我们可使 $a_{11} \neq 0$, 所以二次型可以写成 $a_{11}(\Sigma x_i b_{ij} x_j)$, 这里 $b_{ij} = \frac{a_{ij}}{a_{11}}, b_{11} = 1$. 由于矩阵的对称性, 含有 x_1 的项为

$$x_1^2 + 2\sum_{j=2}^{n} b_{1j}x_1x_j = \left(x_1 + \sum_{j=2}^{n} b_{1j}x_j\right)^2 - \left(\sum_{j=2}^{n} b_{1j}x_j\right)^2.$$

这个“完全平方”的构成暗示着一个变换

$$y_1 = x_1 + \sum_{j=2}^{n} b_{1j}x_j, \quad y_2 = x_2, \quad \cdots, \quad y_n = x_n.$$

那么 y_1 只出现在 y_1^2 项中. 原来的二次型现在就变为 $a_{11}y_1^2 + \Sigma y_j c_{jk} y_k$, 这里的下标 j 和 k 是从 2 跑到 n. 剩下的部分是 $n-1$ 个变量 $y_2, \cdots, y_n$ 的二次型, 对这个二次型使用同样的方法. 这个方法可以重复进行下去 (归纳论证) 直到剩下的二次型的新系数全都是零为止. 因此我们有

定理 15 通过变量的非奇异线性变换, 满足条件 $1+1 \neq 0$ 的任意域上的二次型可以化为对角二次型

$$d_1y_1^2 + d_2y_2^2 + \cdots + d_ry_r^2, \text{ 每个 } d_i \neq 0. \tag{32}$$

非零对角线元素的个数 r 是一个不变量.

这个数 r 称为已知二次型 $\boldsymbol{XAX}^{\mathrm{T}}$ 的秩. 因为 r 是化简后的二次型 (32) 的对角矩阵 $\boldsymbol{D}$ 的秩, 所以它的不变性是显然的. 这个秩必等于原来二次型的矩阵 $\boldsymbol{A}$ 的秩, 因为根据定理 14, 我们的变换把 $\boldsymbol{A}$ 化为 $\boldsymbol{D} = \boldsymbol{PAP}^{\mathrm{T}}$, 而我们已经知道 (8.9 节定理 19), 秩在更一般的变换 $\boldsymbol{A} \longmapsto \boldsymbol{PAQ}$ 之下是一个不变量.

如果 n 个变量的二次型 $\boldsymbol{XAX}^{\mathrm{T}}$ 的秩是 n, 那么称它是非奇异的, 因为这意味着矩阵 $\boldsymbol{A}$ 是非奇异的.

在对角二次型 (32) 中, 秩 r 是不变量, 而系数并不是不变量, 因为用不同的方法化简二次型可以产生不同的系数组 $d_1, \cdots, d_r$. 下一节我们将得到实数域这个特殊情形下的全系不变量.

习 题

1. 在有理数域上, 把 9.7 节习题 4 中的每个二次型化为对角型.
2. 在模 5 整数域 $\mathbf{Z}_5$ 上, 把 $2x^2 + xy + 3y^2$ 化成对角型.
3. 证明: 在域 $\mathbf{Z}_5$ 上, 每个二次型可以通过线性变换化成形式 $\Sigma d_i y_i^2$, 其中每个系数 $d_i = 0, 1$ 或 2.
4. 证明: 在有理数域上, 二次型 $x_1^2 + x_2^2$ 可以变换成两种不同的对角型: $9y_1^2 + 4y_2^2$ 和 $2z_1^2 + 8z_2^2$.
5. 当矩阵 $\boldsymbol{A}$ 是下列情形时, 求出矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{PAP}^{\mathrm{T}}$ 是对角矩阵.

(a) $\boldsymbol{A} = \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix}$, (b) $\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$, (c) $\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 0 \end{pmatrix}$.

6. 求出所有的把实二次型 $x_1^2+\cdots+x_n^2$ 化为 $y_1^2+\cdots+y_n^2$ 的线性变换.
7. 严格证明：二次型 xy 在群 $L_2(\mathbf{Z}_2)$ 之下不等价于对角型.

9.9 全线性群之下的实二次型

在解析几何中, 二次曲线和二次曲面都是用实二次多项式函数来描述. 在实数域上, 对角型 (32) 的每一项可以通过变量代换 $y_i'=|d_i|^{\frac{1}{2}}y_i$ 进一步化简, 使得项 $d_iy_i^2$ 变成 $\pm y_i'^2$. 对所有变量同时实行代换, 就把二次型化为形式 $\Sigma\pm y_i^2$. 在这个和中, 可以置换变量, 使得正的平方项都在前面. 这就证明了

定理 16 实数域上任意二次型可以通过变量的非奇异线性变换化成如下形式：

$$z_1^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_r^2. \tag{33}$$

定理 17 出现在简化形式 (33) 中正平方项的个数 p 是给定二次型 Q 的不变量, 也就是说, p 只依赖于二次型而不依赖于化简的方法 (西耳维斯特 (Sylvester) 惯性律).

证明 假设存在另外的简化形式

$$y_1^2+\cdots+y_q^2-y_{q+1}^2-\cdots-y_r^2, \tag{34}$$

它含有 q 个正项. 因为这两个简化形式是从同一个 Q 通过非奇异变换得到的, 所以存在一个非奇异变换把 (33) 化成 (34). 我们可以把这个变换方程看成坐标变换 (图像固定坐标移动的变换), 那么 (33) 和 (34) 表示固定向量 $\boldsymbol{\xi}$ 的同一个二次函数 $Q(\boldsymbol{\xi})$, 而 $\boldsymbol{\xi}$ 相对于一组基的坐标是 $z_1,\cdots,z_r$, 相对于另一组基的坐标是 $y_1,\cdots,y_r$.

假定 $q<p$. 如果公式 (33) 中 $z_{p+1}=\cdots=z_r=0$, 则 $Q(\boldsymbol{\xi})\geqslant 0$. 满足这 $r-p$ 个方程的全体向量 $\boldsymbol{\xi}$ 构成 $n-(r-p)$ 维子空间 S_1(在这个子空间中存在 $n-(r-p)$ 个坐标 $z_1,\cdots,z_p,z_{r+1},\cdots,z_n$). 类似地, (34) 式中, 如果每个 $\boldsymbol{\xi}\neq 0$ 的坐标满足 $y_1=\cdots=y_q=y_{r+1}=\cdots=y_n=0$, 则使得 $Q(\boldsymbol{\xi})<0$. 这些条件确定一个 $r-q$ 维子空间 S_2. S_1 和 S_2 两个子空间的维数之和是

$$n-(r-p)+(r-q)=n+(p-q)>n$$

因此 S_1 和 S_2 具有公共的非零向量 $\boldsymbol{\xi}$, 因为由 7.8 节定理 17 可知交 $S_1\cap S_2$ 的维数是正的. 对于这个公共向量 $\boldsymbol{\xi}$, 由 (33) 有 $Q(\boldsymbol{\xi})\geqslant 0$, 而由 (34) 有 $Q(\boldsymbol{\xi})<0$, 显然得出矛盾. 如果假定 $q>p$, 则将导出类似的矛盾, 所以 $q=p$, 定理证完.

这个结果表明, 任意实二次型可以通过线性变换化为一种且只化为一种 (33) 那样的二次型. 所以这种类型的表达式 $\Sigma\pm z^2$ 是实二次型在全线性群之下的标准型. 这个标准型本身由所谓*符号差*$\{+,\cdots,+,-,\cdots,-\}$ 唯一确定, 这个符号差是

由 p 个正号, $r-p$ 个负号组成, 其中 r 是二次型的秩. 这个符号差是通过 r 和 $s=p-(r-p)=2p-r$ 确定的 (s 是正号个数减去负号个数). 有时称这个整数 s 为二次型的*符号差*. r 和 s 一起构成全系 (数值) 不变量, 这是根据两个二次型等价当且仅当它们可化成同一个标准型 (33).

定理 18 *两个实二次型在全线性群之下等价当且仅当它们有相同的秩和符号差.*

n 个变量的实二次型 $Q=\boldsymbol{XAX}^{\mathrm{T}}$, 如果当 $\boldsymbol{X}\neq 0$ 时可推出 $Q>0$, 那么称 Q 是*正定二次型*; 在同样条件下实对称矩阵$\boldsymbol{A}$ 称为*正定矩阵*. 如果我们考虑标准型 (33), 显然, Q 是正定二次型当且仅当标准型是 $z_1^2+\cdots+z_n^2$. 这是因为, n 个平方和除了所有项都是零之外总是正的, 而且当取 $\boldsymbol{X}$ 是第 n 个单位向量 ε_n 时, (33) 中的 $\boldsymbol{XAX}^{\mathrm{T}}\leqslant 0$, 除非 $p=n$, 于是, 我们证明了

定理 19 *实二次型是正定的当且仅当它的标准型是* $z_1^2+\cdots+z_n^2$.

根据定理 14, 这意味着 $\boldsymbol{A}=\boldsymbol{PIP}^{\mathrm{T}}$, 它给出下面进一步的结果.

定理 20 *实对称矩阵 $\boldsymbol{A}$ 是正定的当且仅当存在一个实非奇异矩阵 $\boldsymbol{P}$ 使得* $\boldsymbol{A}=\boldsymbol{PP}^{\mathrm{T}}$.

二次型 $\boldsymbol{XAX}^{\mathrm{T}}$ 确定了在 n 维向量空间中的一条轨迹, 它是由满足 $\boldsymbol{XAX}^{\mathrm{T}}=1$ 的所有点 $\boldsymbol{X}$ 组成. 标准型 (33) 意味着, 通过适当的非奇异线性变换可把轨迹化为具有方程

$$z_1^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_r^2=1$$

的曲线, 例如, 在平面上, 秩为 2 的简化了的方程是

$$x^2+y^2=1,\quad x^2-y^2=1,\quad -x^2-y^2=1.$$

它们分别表示圆、等轴双曲线或根本没有轨迹. 秩为 0 的唯一的二次型是 $0=1$; 秩为 1 的二次型是 $x^2=1$(它表示两条直线 $x=\pm1$) 或者 $-x^2=1$(没有轨迹). 在 8.8 节中我们证明了 (定理 15 的推论 2), 平面上任意非奇异线性变换可以表示为切变换、压缩 (或伸长) 与反射的乘积. 因此方程 $ax^2+bxy+cy^2=1$ 表示的任意有心二次曲线可以相继通过切变换、压缩与反射化为我们在上面列举出的几种形式之一. 在几何上, 这些结果是合理的: 椭圆可以沿着一根轴压缩成圆; 但是显然找不到一系列线性变换可以把圆 $x^2+y^2=1$ 化成等轴双曲线 $x^2-y^2=1$. 这就是平面情形符号差的不变性的几何意义.

在研究两个变量的函数的极大值和极小值时, 符号差是有用的. 设 $z=f(x,y)$ 是一个光滑函数, 它的一阶偏导数 f_x 和 f_y 在 $x=x_0$, $y=y_0$ 处都为零, 因此 z 按 $h=x-x_0, k=y-y_0$ 的幂的泰勒级数展开式中, 没有一次项. 这个展开式是

$$f(x_0+h,y_0+k)=f(x_0,y_0)+\frac{1}{2}[ah^2+2bhk+ck^2]+\cdots,$$

其中系数 a, b, c 是偏导数

$$a = f_{xx}(x_0, y_0), \quad b = f_{xy}(x_0, y_0), \quad c = f_{yy}(x_0, y_0).$$

当 h 和 k 的值很小时, 起支配作用的是方括号中的那一项, 它是变量 h 和 k 的实系数二次型. 如果这个二次型的秩是 2, 那么它可以利用变换后的变量 h' 和 k' 表示成 $\pm h'^2 \pm k'^2$. 如果两项的符号都是加号, 那么 (x_0, y_0) 点附近的函数值 $f(x_0+h, y_0+k)$ 总是超过 $f(x_0, y_0)$, 所以 z 有相对*极小值*. 如果两项的符号都是减号, 那么 z 有相对*极大值*. 如果一项符号是加号另一项符号是减号, 则二次型可以取正值也可以取负值, 所以 (x_0, y_0) 既不是极大点也不是极小点, 而是*鞍点*(像马鞍, 或者像两个山峰之间的隘口, 高度 z 沿一个方向是增加的, 而沿另一个方向是减少的). 因此, 极大点、极小点和鞍点是根据二次型的符号差来区分的. 三个变量或更多变量的函数的*临界点*有类似的结果.

习　题

1. 证明：实二次函数 $ax^2 + bxy + cy^2$ 是正定的当且仅当 $a > 0$ 且 $4ac - b^2 > 0$.
2. 证明：正定对称矩阵的主对角线上的元素都是正的.
3. 把下列各实二次型化为定理 16 所述的标准型, 并求出每个二次型的秩和符号差：
   (a) $9x_1^2 + 12x_1x_2 + 79x_2^2$,　　(b) $2x_1^2 - 12x_1x_2 + 18x_2^2$,
   (c) $-2x_1^2 - 4x_1x_2 + 22x_2^2 + 12x_2x_3 + 6x_3x_1 - x_3^2$.
4. 描述三维空间中实二次型的各种可能标准型的几何轨迹.
5. 证明：复系数齐次二次型在全复线性群之下总是与平方和 $z_1^2 + \cdots + z_r^2$ 等价.
6. 证明：n 个变量的复系数的两个二次型在全线性群之下等价当且仅当它们有相同的秩.
7. 证明：双线性函数 $\boldsymbol{XAY}^{\mathrm{T}}$ 是 “内积” 当且仅当 $\boldsymbol{A}$ 是对称的并且是正定的.
8. 如果二次型的秩等于符号差, 那么称这个二次型是*半正定的*, 对这种二次型叙述并证明类似于定理 19 的命题.
9. 对半正定二次型, 叙述并证明类似于定理 20 的命题.
10. (a) 列出四个变量非奇异二次型的所有类型.
    (b) 至少对两个二次型, 几何地描述一下它们在 $\mathbf{R}^4$ 中相应的轨迹.

9.10　正交群之下的二次型

实二次型在正交变换之下可以化简成什么样呢? 一个正交变换 $\boldsymbol{Y} = \boldsymbol{XP}$ 把 $\boldsymbol{XAX}^{\mathrm{T}}$ 变换成 $\boldsymbol{Y}(\boldsymbol{P}^{-1}\boldsymbol{A}(\boldsymbol{P}^{-1})^{\mathrm{T}})\boldsymbol{Y}^{\mathrm{T}}$. 因为 $\boldsymbol{P}$ 是正交矩阵, 所以新矩阵可以写成[①] $\boldsymbol{P}^{-1}\boldsymbol{A}(\boldsymbol{P}^{-1})^{\mathrm{T}} = \boldsymbol{P}^{-1}\boldsymbol{AP}$.

① 两个对称矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{P}^{-1}\boldsymbol{AP}$($\boldsymbol{P}$ 是正交矩阵) 有时称为是*正交相合的*.

在平面上, 椭圆在正交变换 (旋转或反射) 之下决不能产生圆; 我们至多可以使椭圆的轴旋转到标准位置上. 长轴可以看作最长的直径. 为重述这个最大值性质, 考虑任意实二次函数 $Q(\boldsymbol{\xi}) = ax^2 + cy^2$, 其中 $a \leqslant c$, 并且没有 xy 项. 那么 $Q(\boldsymbol{\xi}) \leqslant cx^2 + cy^2 = c(x^2 + y^2)$, 这就意味着, 在单位圆 $x^2 + y^2 = 1$ 的所有点上 Q 的最大值是 c, 并且在 $x = 0, y = 1$ 点上取这个最大值. 反过来, 下面的引理保证 Q 中没有 xy 项.

引理 *如果实二次函数 $Q = ax^2 + 2bxy + cy^2$ 在单位圆 $x^2 + y^2 = 1$ 的所有点中有一个最大值, 并且在点 $x = 0, y = 1$ 上取得, 那么 $b = 0$.*

证明 把 Q 看作一个变量 x 的 (双值) 函数, 这里 y 同 x 的关系隐含在等式 $x^2+y^2 = 1$ 中. 两边求微商, 我们得到 $2x + 2y\dfrac{\mathrm{d}y}{\mathrm{d}x} = 0$, 所以导数 $y' = \dfrac{\mathrm{d}y}{\mathrm{d}x} = -\dfrac{x}{y}$. Q 的导数是

$$Q' = (ax^2 + 2bxy + cy^2)' = 2ax + 2by + 2bxy' + 2cyy'.$$

置 $x = 0, y = 1$, 求出 y' 的值后一起代入上式, 我们得到 $Q' = 2b$. 但是 Q 在 $x = 0, y = 1$ 点处达到最大值, 所以这个导数一定为零, 因此 $2b = 0$. 证毕

现在回到 n 个变量的二次型. 在 n 维空间中, 单位超球面 $\Sigma x_i^2 = 1$ 是封闭的有界集合, 它上面的点都是长度为 1 的向量. 在这个超球面上, 实二次型 $Q(\boldsymbol{\xi}) = \sum\limits_{i,j} x_i a_{ij} x_j$所取的值有一个上界 $\sum\limits_{i,j} |a_{ij}|$. 因为 $Q(\boldsymbol{\xi})$ 是 $\boldsymbol{\xi}$ 的连续函数, 所以 $Q(\boldsymbol{\xi})$ 在 S 上有最大值① λ_1. 换句话说, 在所有单位长度的向量 $\boldsymbol{\xi}$ 中间, 存在一个向量 $\boldsymbol{\xi}_0$, 在 $\boldsymbol{\xi}_0$ 上, $Q(\boldsymbol{\xi})$ 取它的最大值 λ_1. 因为 $\boldsymbol{\xi}_0$ 的长度为 1, 所以我们可以选取 $\boldsymbol{\alpha}_1 = \boldsymbol{\xi}_0$ 作为新的标准正交基 $\alpha_1, \cdots, \alpha_n$ 的第一个向量 (7.11 节的定理 21). 设 $\boldsymbol{\xi}$ 对于这组基的新坐标是 $y_1, \cdots y_n$, 那么二次型按照新坐标可表示为 $Q(\boldsymbol{\xi}) = \Sigma y_i b_{ij} y_j$, 其中系数 b_{ij} 组成一个新矩阵. Q 的最大值 λ_1 通过坐标为 $(1, 0, \cdots, 0)$ 的向量 $\boldsymbol{\alpha}_1$ 给出, 所以代入 $\boldsymbol{\alpha}_1$ 可知最大值 λ_1 就等于 b_{11}. 如果我们进一步限制变量, 除了 y_1 和 y_i 两个变量之外其他都是零, 那么这个最大值仍然保持不变. 因此 $y_1 = 1, y_i = 0$ 是二次型 $b_{11}y_1^2 + 2b_{1i}y_1y_i + b_{ii}y_i^2$ 在条件 $y_1^2 + y_i^2 = 1$ 之下的极大点. 那么引理 (用 y_i 代替 x) 断言, 交叉乘积的系数 b_{1i} 是零, 这种论证应用于 $i = 2, 3, \cdots, n$ 每一种情形. 因此, Q 按照这些坐标 $y_1, \cdots, y_n$ 来表示时, 去掉了所有 y_i 与 y_1 交叉乘积项, 于是变成

$$Q(\boldsymbol{\xi}) = \lambda_1 y_1^2 + \sum_{i=2}^{n}\sum_{j=2}^{n} y_i b_{ij} y_j, \quad \boldsymbol{B} = (b_{ij}) = \boldsymbol{B}^{\mathrm{T}}. \tag{35}$$

第一个系数 λ_1 不是向量而是标量 (是在球面 $|\boldsymbol{\xi}| = 1$ 上 $Q(\boldsymbol{\xi})$ 的最大值).

① 同微积分中一样, 这里我们假定下述事实成立: 在有界闭集上的连续函数在这个集合上有一个最大值.

在 (35) 式中差 $Q^*(\boldsymbol{\xi})=Q(\boldsymbol{\xi})-\lambda_1 y_1^2$ 是 $n-1$ 个变量 $y_2,\cdots,y_n$ 的二次型. 这些变量是由 $n-1$ 个新基向量 $\alpha_2,\cdots,\alpha_n$ 张成的空间 S_{n-1} 中的坐标. 在这个空间 (它是第一个基向量 $\boldsymbol{\xi}_0$ 的正交补) 中, 我们可以重复应用同样的方法, 选择新的标准正交基使得 $Q^*(\boldsymbol{\xi})$ 在 $|\boldsymbol{\xi}|=1$ 上取最大值; 这就从二次型中分出另外一个对角线元素. 最后我们求出一组主轴基, 对于这组基有

$$Q(\boldsymbol{\xi})=\lambda_1 z_1^2+\lambda_2 z_2^2+\cdots+\lambda_n z_n^2,\quad \lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n. \tag{36}$$

这里 $z_1,\cdots,z_n$ 是 $\boldsymbol{\xi}$ 相对于 $\alpha_1,\beta_2,\gamma_3,\cdots$ 这组基的坐标, 而 $\alpha_1,\beta_2,\gamma_3,\cdots$ 是在逐次满足最大要求时选取的. 第一个向量 α_1 使 $Q(\boldsymbol{\xi})$ 达到最大值 λ_1, 只在条件 $|\boldsymbol{\xi}|=1$ 之下. 第二个基向量 β_2 选为在这个空间中与 $\boldsymbol{\alpha}_1$ 正交的向量; 这也就是, $\boldsymbol{\eta}=\boldsymbol{\beta}_2$ 是适合 $|\boldsymbol{\eta}|=1,(\boldsymbol{\eta},\boldsymbol{\alpha}_1)=0$ 的所有向量 $\boldsymbol{\eta}$ 中间, 使 $Q(\boldsymbol{\eta})$ 达到最大值 λ_2 的一个向量. 第三个基向量 γ_3 是适合 $|\boldsymbol{\zeta}|=1$, 并与 $\boldsymbol{\alpha}_1,\boldsymbol{\beta}_2$ 正交的所有向量 $\boldsymbol{\zeta}$ 中间, 使 $Q(\boldsymbol{\zeta})$ 达到最大值的一个向量, 等等. 这个逐次最大问题可以通过具有三个不等轴 $a>b>c>0$ 的椭球 (以倒过来的形式, 即极大换作极小) 加以形象地描述. 最短主轴 c 是最小直径; 次主轴 b 是所有同最短主轴垂直的直径中间最短的一个, 等等.

于是 (36) 式的系数 $\lambda_1,\cdots,\lambda_n$ 可表征为某一极值问题的解, 这个极值问题只依赖于 Q, 而不依赖于特殊的坐标系. 在化简过程中, 只有当第一个最大值 (或者后面的某一最大值) 是由两个或更多个长度为 1 的不同向量 $\boldsymbol{\xi}_0$ 和 $\boldsymbol{\eta}_0$ 给出的时候才可能产生极值不确定的情况. 即使在这种情况, 我们仍可证明 λ_i 是唯一的 (10.4 节).

这就证明了下面的主轴定理.

定理 21 *任意 n 个变量的实二次型, 对于适当的标准正交基可以取为对角型* (36).

由定理 1, 这组新基 $\boldsymbol{\alpha}_1^*,\cdots,\boldsymbol{\alpha}_n^*$ 可以按照原基 $\boldsymbol{\varepsilon}_1=(1,0,\cdots,0),\cdots,\varepsilon_n=(0,\cdots,0,1)$ 表示成 $\boldsymbol{\alpha}_i^*=\sum\limits_j p_{ij}\boldsymbol{\varepsilon}_j$. 进一步, 因为向量 $\boldsymbol{\alpha}_1^*,\cdots,\boldsymbol{\alpha}_n^*$ 是标准的而且是正交的, 所以系数矩阵 $\boldsymbol{P}=(p_{ij})$ 是正交矩阵, 像定理 2 那样, 老坐标 $x_1,\cdots,x_n$ 则可按照新坐标 $x_1^*,\cdots,x_n^*$ 表示成 $x_j\sum\limits_i x_i^* p_{ij}$; 换句话说, 我们已对二次型做了变量的正交变换. 于是定理 21 的 "坐标变换" 形式的结果可改写成 "点变换" 的形式, 即

推论 1 *n 个变量的任意齐次二次函数可以通过正交点变换化为对角型* (36).

这两个结果都称为 "主轴定理". 如果二次型用与它对应的矩阵来代替, 那么这个定理断言

推论 2 *对任意实对称矩阵 $\boldsymbol{A}$, 存在实正交矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{PAP}^{\mathrm{T}}=\boldsymbol{PAP}^{-1}$ 是对角矩阵.*

换句话说, 我们证明了任意实对称矩阵与对角矩阵相似. 同定理 4 相比较, 我们看出标准型 (36) 中的 $\lambda_1,\cdots,\lambda_n$ 恰好是矩阵 $\boldsymbol{A}$ 的全体特征值.

在平面上，方程 $Q(\boldsymbol{\xi})=1$ 的标准型只不过是 $\lambda_1x^2+\lambda_2y^2=1$. 这类方程包括椭圆 $(\lambda_1\geqslant\lambda_2>0)$ 或双曲线 $(\lambda_1>0>\lambda_2)$ 的普通的标准方程，这些系数确定了轴的长度. 在三维空间中，对三个系数 $\lambda_1,\lambda_2,\lambda_3$ 有类似的解释. 如果这三个系数都是正的，则轨迹 $Q=1$ 是一个椭球；如果其中一个是负的，则是单叶双曲面；如果其中两个是负的，则是双叶双曲面；如果三个系数都是负的，则根本没有轨迹. (注意符号差和秩在这里所起的作用.)

同定理 4 的推论相比较，我们看出 (对于对称矩阵 $\boldsymbol{A}$) 二次函数 $\boldsymbol{XAX}^{\mathrm{T}}$ 的主轴正是线性变换 $\boldsymbol{X}\longmapsto\boldsymbol{XA}$ 的特征向量. 因此得到

推论 3 对于实对称矩阵$\boldsymbol{A}$，线性变换 $\boldsymbol{X}\longmapsto\boldsymbol{XA}$ 有一组基是由具有实特征值的正交特征向量构成.

推论 4 每个非奇异实矩阵 $\boldsymbol{A}$ 可以表示成乘积 $\boldsymbol{A}=\boldsymbol{SR}$，其中 $\boldsymbol{S}$ 是对称正定矩阵，$\boldsymbol{R}$ 是正交矩阵.

证明 我们已经知道 (定理 20)$\boldsymbol{AA}^{\mathrm{T}}$ 是对称的，并且是正定的，根据主轴定理，存在正交矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{AA}^{\mathrm{T}}\boldsymbol{P}$ 是对角矩阵，并且是正定的. 于是对角线元素都是正的，通过求这些元素的平方根，我们得到一个正定对角矩阵 $\boldsymbol{T}$，满足 $\boldsymbol{T}^2=\boldsymbol{P}^{-1}\boldsymbol{AA}^{\mathrm{T}}\boldsymbol{P}$，因此正定对称矩阵 $\boldsymbol{S}=\boldsymbol{PTP}^{-1}$ 满足 $\boldsymbol{S}^2=\boldsymbol{AA}^{\mathrm{T}}$. 如果我们证明了 $\boldsymbol{R}=\boldsymbol{S}^{-1}\boldsymbol{A}$ 是正交的，则有 $\boldsymbol{A}=\boldsymbol{SR}$，如推论所要求的，于是推论得证. 而实际上 $\boldsymbol{RR}^{\mathrm{T}}=\boldsymbol{S}^{-1}\boldsymbol{AA}^{\mathrm{T}}(\boldsymbol{S}^{-1})^{\mathrm{T}}=\boldsymbol{S}^{-1}\boldsymbol{S}^2(\boldsymbol{S}^{-1})^{\mathrm{T}}=\boldsymbol{S}(\boldsymbol{S}^{-1})^{\mathrm{T}}=\boldsymbol{SS}^{-1}=\boldsymbol{I}$. 这是因为对称矩阵 $\boldsymbol{S}$ 有性质 $(\boldsymbol{S}^{-1})^{\mathrm{T}}=\boldsymbol{S}^{-1}$.

推论 5 设 $\boldsymbol{A}$ 是任意实对称矩阵，$\boldsymbol{B}$ 是任意正定 (实) 对称矩阵，那么存在一个实非奇异矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{PAP}^{-1}$ 和 $\boldsymbol{PBP}^{-1}$ 同时为对角矩阵.

我们把这个推论的证明留作习题. 求满足 $\boldsymbol{A\xi}_j=\lambda_j\boldsymbol{B\xi}_j$ 的一组向量 $\boldsymbol{\xi}_j$，称为广义特征向量问题，它的解在振动理论中起着基本作用.

习 题

1. 考虑实二次型 $ax^2+2bxy+cy^2$.
   (a) 证明：在正交变换下 $a+c$ 和 b^2-ac 是不变量.
   (b) 设 $\cot 2\alpha=\dfrac{a-c}{2b}$，证明：这个二次型可以通过正交变换

$$x=x'\cos\alpha-y'\sin\alpha,\ y=x'\sin\alpha+y'\cos\alpha$$

   化成对角型.
2. 证明：每个实斜对称矩阵 $\boldsymbol{A}$ 具有形式 $\boldsymbol{A}=\boldsymbol{P}^{-1}\boldsymbol{BP}$，其中 $\boldsymbol{P}$ 是正交矩阵，$\boldsymbol{B}^2$ 是对角矩阵.
3. 按照已给出的方法，通过正交变换把下列二次型化为对角型：
   (a) $5x^2-6xy+5y^2$， (b) $2x^2+4\sqrt{3}xy-2y^2$.

4. 把正交变换

$$3x_1 = 2y_1 - y_2 + 2y_3, \quad 3x_2 = -y_1 + 2y_2 + 2y_3, \quad 3x_3 = 2y_1 + 2y_2 - y_3$$

作用到二次型 $9x_1^2 - 9x_2^2 + 18x_3^2$ 上, 其结果 Q 是 y_1, y_2, y_3 的二次型, 直接证明: 向量 $\left(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}\right)$ 给出 Q 在 $y_1^2 + y_2^2 + y_3^2 = 1$ 上的最大值为 18. 用微积分中的方法来检验.

5. 在单位圆 $x = \cos\theta, y = \sin\theta$ 上考虑二次型 $ax^2 + 2bxy + cy^2$. 证明它的极值是 (参看习题 1):

$$\frac{(a+c) \pm \sqrt{(a+c)^2 - 4\Delta}}{2}, \quad \Delta = ac - b^2.$$

6. 证明: 找不到含有有理元素的正交矩阵, 把 xy 化为对角型.
7. 保持 $x_1^2 + x_2^2 + x_3^2 - x_4^2$ 不变的线性变换称为罗伦兹 (Lorentz) 变换. 证明: 矩阵 $\boldsymbol{P}$ 定义一个罗伦兹变换当且仅当 $\boldsymbol{P}^{-1} = \boldsymbol{S}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S} = \boldsymbol{S}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}^{-1}$, 其中 $\boldsymbol{S}$ 是个特殊的对角矩阵, 其对角线元素为 $1, 1, 1, -1$.
8. (a) 证明: 如果 $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{R}$, 其中 $\boldsymbol{S}$ 是对称矩阵, $\boldsymbol{R}$ 是正交矩阵, 那么 $\boldsymbol{S}^2 = \boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$.
   *(b) 证明: 只存在一个正定矩阵 $\boldsymbol{S}$, 它满足 $\boldsymbol{S}^2 = \boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$. (提示: $\boldsymbol{S}^2$ 的任意特征向量一定是 $\boldsymbol{S}$ 的一个特征向量.)

*9. 证明定理 21 的推论 5. (提示: 把 $\boldsymbol{X}\boldsymbol{A}\boldsymbol{X}^{\mathrm{T}}$ 看作具有内积 $\boldsymbol{X}\boldsymbol{B}\boldsymbol{X}^{\mathrm{T}}$ 的欧几里得向量空间的二次函数, 再根据定理 20, 把 $\boldsymbol{B}$ 写成 $\boldsymbol{B} = \boldsymbol{P}\boldsymbol{P}^{\mathrm{T}}$.)

9.11 仿射群和欧几里得群之下的二次型

下面考虑含有坐标 $x_1, \cdots, x_n$ 的向量 $\boldsymbol{\xi}$ 的任意*非齐次二次函数*

$$f(\xi) = \sum_i \sum_j x_i a_{ij} x_j + \sum_k b_k x_k + c \quad (i, j, k = 1, \cdots, n). \tag{37}$$

这可以写成 $f(\boldsymbol{\xi}) = \boldsymbol{X}\boldsymbol{A}\boldsymbol{X}^{\mathrm{T}} + \boldsymbol{B}\boldsymbol{X}^{\mathrm{T}} + c$, 其中 $\boldsymbol{A} = (a_{ij})$ 是对称矩阵, $\boldsymbol{B} = (b_1, \cdots, b_n)$ 是行矩阵. 在单变量函数 $f = ax^2 + bx + c$ 的简单情形中, 我们看出, 平移 $x = y + k$ 使二次项系数 a 保持不变, 这是因为

$$f = a(y+k)^2 + b(y+k) + c = ay^2 + (2ak+b)y + ak^2 + bk + c. \tag{38}$$

对 n 个变量的情形做类似的计算, 平移 $\boldsymbol{X} \longmapsto \boldsymbol{Y} = \boldsymbol{X} - \boldsymbol{K}$($\boldsymbol{K}$ 是行矩阵) 给出

$$\begin{aligned} f(\boldsymbol{\xi}) &= (\boldsymbol{Y}+\boldsymbol{K})\boldsymbol{A}(\boldsymbol{Y}+\boldsymbol{K})^{\mathrm{T}} + \boldsymbol{B}(\boldsymbol{Y}+\boldsymbol{K})^{\mathrm{T}} + c \\ &= \boldsymbol{Y}\boldsymbol{A}\boldsymbol{Y}^{\mathrm{T}} + \boldsymbol{K}\boldsymbol{A}\boldsymbol{Y}^{\mathrm{T}} + \boldsymbol{Y}\boldsymbol{A}\boldsymbol{K}^{\mathrm{T}} + \boldsymbol{K}\boldsymbol{A}\boldsymbol{K}^{\mathrm{T}} + \boldsymbol{B}\boldsymbol{Y}^{\mathrm{T}} + \boldsymbol{B}\boldsymbol{K}^{\mathrm{T}} + c. \end{aligned}$$

因为乘积 $\boldsymbol{Y}\boldsymbol{A}\boldsymbol{K}^{\mathrm{T}}$(行矩阵 × 矩阵 × 列矩阵) 是一个标量, 所以它等于它的转置 $\boldsymbol{K}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Y}^{\mathrm{T}} = \boldsymbol{K}\boldsymbol{A}\boldsymbol{Y}^{\mathrm{T}}$, 总起来有

$$f(\boldsymbol{\xi})=\boldsymbol{YAY}^{\mathrm{T}}+(2\boldsymbol{KA}+\boldsymbol{B})\boldsymbol{Y}^{\mathrm{T}}+\boldsymbol{KAK}^{\mathrm{T}}+\boldsymbol{BK}^{\mathrm{T}}+c. \tag{39}$$

这确实与公式 (38) 类似. 这就证明了

引理 *平移使二次函数 $f(\boldsymbol{\xi})$ 的齐次二次项部分对应的矩阵 $\boldsymbol{A}$ 保持不变.*

另一方面, 齐次线性变换 $\boldsymbol{Y}=\boldsymbol{YP}$ 把 $f(\boldsymbol{\xi})$ 变到 $\boldsymbol{Y}(\boldsymbol{PAP}^{\mathrm{T}})\boldsymbol{Y}^{\mathrm{T}}+(\boldsymbol{BP}^{\mathrm{T}})\boldsymbol{Y}^{\mathrm{T}}+c$. 在这个二次函数中, 二次项的新矩阵是 $\boldsymbol{PAP}^{\mathrm{T}}$, 这恰好同单独是齐次二次型的变换情形一样.

现在通过具有方程 $\boldsymbol{X}=\boldsymbol{YP}+\boldsymbol{K}$($\boldsymbol{P}$ 是正交矩阵) 的刚体运动来化简实函数 $f(\boldsymbol{\xi})$. 根据上述说明, 只是 $\boldsymbol{P}$ 所对应的正交变换可以化简二次项的矩阵 $\boldsymbol{A}$, 确切地说, 是齐次二次型对应的矩阵 $\boldsymbol{A}$. 像在 9.10 节中那样, 我们得到 (用新的系数 b_i')

$$f(\boldsymbol{\xi})=\lambda_1 z_1^2+\cdots+\lambda_n z_n^2+b_1' z_1+\cdots+b_n' z_n+c.$$

同非零的 λ_j 相联系的 b_j', 现在可以通过简单的 "配方" 法用平移 $y_j=z_j+\dfrac{b_j'}{2\lambda_j}$ 消去. 我们把变量置换一下, 使得非零的 λ 项都移到前面, 于是我们得到

$$f(\boldsymbol{\xi})=\lambda_1 y_1^2+\cdots+\lambda_r y_r^2+b_{r+1}' z_{r+1}+\cdots+b_n' z_n+c'.$$

如果这个函数的线性部分不正好是常数 c', 那么可以通过适当的平移和正交变换, 像在定理 13 中那样, 化为 dy_{r+1} 的形式. 这个变换并不影响前 r 个变量. 最后结果是下面形式中的一种

$$f(\boldsymbol{\xi})=\lambda_1 y_1^2+\cdots+\lambda_r y_r^2+dy_{r+1}, \tag{40}$$

$$f(\boldsymbol{\xi})=\lambda_1 y_1^2+\cdots+\lambda_r y_r^2+c'. \tag{41}$$

这里 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_r$, 所有 λ_i 都不为零, $d>0$.

定理 22 *在所有刚体运动的欧几里得群之下, 任意实二次型 (37) 等价于 (40) 或 (41) 式中的一种.*

这些简化形式确实是实二次型在欧几里得群之下的标准型, 但是证明比较困难, 只概括叙述如下.

这些 λ_i 是 (37) 式相对应的矩阵 $\boldsymbol{A}$ 的全部特征值 (见 9.10 节); 它们的唯一性 (包括重数) 将在 10.4 节中证明. 特别是, (40) 或 (41) 式中的平方项个数 r 是一个不变量; 注意, r 也是 $\boldsymbol{A}$ 的秩, 它在变换 $\boldsymbol{A}\longmapsto\boldsymbol{PAP}^{\mathrm{T}}$ 之下是不变的. 利用微积分, 可以对不变量 d 和 c' 作更为简单的直观描述. 我们考虑使得向量

$$\operatorname{grad} f=\left(\frac{\partial f}{\partial y_1},\cdots,\frac{\partial f}{\partial y_n}\right)$$

是零向量的轨迹. 在 (41) 式的情况下, 这个轨迹是子空间 $y_1 = \cdots = y_r = 0$, 而 c' 是 $f(\boldsymbol{\xi})$ 在这个 (不变) 轨迹上的常数值. 在 (40) 式的情况下, 这个轨迹是空的, 因为 $\dfrac{\partial f}{\partial y_{r+1}} = d \neq 0$; 而 d 可以表征为 $|\mathrm{grad} f|$ 的最小值; 还可以证明, 这个最小值在欧几里得群之下是不变量.

对于仿射变换 $\boldsymbol{X} = \boldsymbol{YP} + \boldsymbol{K}$(其中 $\boldsymbol{P}$ 是非奇异的), 可以采用类似的处理方法. 在化简二次项部分为对角型时, 像在 9.9 节那样所有系数都是 ± 1. 而线性部分可像 9.6 节中那样去处理.

定理 23 *n 个变量的任意实二次函数可以通过仿射变换 (或者通过坐标仿射变换) 化为下面形式中的一种.*

$$y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2 + c \quad (r \leqslant n), \tag{42}$$

$$y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2 + y_{r+1} \quad (r < n). \tag{43}$$

因为平移不影响二次项, 所以根据惯性律 (定理 17), 秩 r 和正项个数 p 一定是不变量.

从几何观点来看, 每个二次函数 $f(\boldsymbol{\xi}) = \boldsymbol{XAX}^{\mathrm{T}} + \boldsymbol{BX}^{\mathrm{T}} + c$ 定义一个*图形*或*轨迹*, 它是由满足方程 $f(\boldsymbol{\xi}) = 0$ 的所有向量 $\boldsymbol{\xi}$ 组成. 在二维空间中, 根据这类二次方程画得的图形就是普通的二次曲线; 在三维空间中, 它是二次曲面; 在一般情形中, 它称为*超二次曲面*(或者二次超曲面). 一个仿射变换 $\boldsymbol{Y} = \boldsymbol{XP} + \boldsymbol{K}$ 作用于这个曲面的方程也就相当于把同一个变换作用到图形的全部点上, 并且我们称新的图形在给定的仿射变换之下*等价于*老的图形.

显然, 上述的在等价之下二次函数分类的结论将给出对应图形的一个类似的分类. 然而, 我们首先看出, 方程 $f(\boldsymbol{\xi}) = 0$ 与同一方程的 "数乘" 积 $cf(\boldsymbol{\xi}) = 0$ 给出同样的轨迹. 这可以用来简化标准型, 例如上面求得的 $y_1^2 - y_2^2 + c = 0$. 当 $c \neq 0$ 时, 这个方程给出的轨迹同 $(c^{-1})y_1^2 - (c^{-1})y_2^2 + 1 = 0$ 给出的轨迹一样; 当 $c > 0$ 时, 通过仿射变换 $y_1 = \sqrt{c}z_1, y_2 = \sqrt{c}z_2$ 可以把它化为 $z_1^2 - z_2^2 + 1 = 0$, 而对于 $c < 0$, 变换 $y_i = \sqrt{-c}z_i$ 给出类似的结果 $z_2^2 - z_1^2 + 1 = 0$. 一般地, 用这个方法总可以把 (43) 式中出现的常数变为 1 或 0. 所以在实数域上的 n 维向量空间中, 任意超二次曲面在仿射群之下等价于下面方程中的一个所给出的轨迹:

$$y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2 + 1 = 0, \tag{44}$$
$$y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2 + y_{r+1} = 0, \tag{45}$$
$$y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2 = 0, \tag{46}$$

这里 $0 \leqslant p \leqslant r \leqslant n$, 在 (45) 的情形中 $r < n$.

在 (44) 式中, 不同的二次型表示仿射不等价的轨迹, 而在 (45) 式中, 变换 $y_{r+1} \longmapsto -y_{r+1}$ 把 p 和 $r-p$ 互换, 于是变换后的二次型与原来的等价.

例如, 平面上满足 $r>0$ 的各种可能类型的轨迹是:

$r=2$		$r=1$	
$x^2+y^2+1=0$	没有轨迹	$\pm x^2+y=0$	抛物线
$x^2-y^2+1=0$	双曲线	$x^2+1=0$	没有轨迹
$-x^2-y^2+1=0$	圆	$-x^2+1=0$	两条平行的直线
$\pm(x^2+y^2)=0$	一点	$x^2=0$	一条直线
$x^2-y^2=0$	两条相交的直线		

特别要注意, 不同的标准函数 x^2+y^2+1 和 x^2+1 给出相同的轨迹 (即, 这个图形一个点也没有). 还有标准函数 x^2+y^2 和 $-x^2-y^2$, x^2+y 和 $-x^2+y$ 也是这样.

习　　题

1. 在欧几里得群之下, 将下列二次型分类:
   (a) $4xz+4y^2+8y+8$,
   (b) $9x^2-4xy+6y^2+3z^2+2\sqrt{5}x+4\sqrt{5}y+12z+16$.
2. 在仿射群之下, 将下列二次型分类:
   (a) $x^2+4y^2+9z^2+4xy+6xz+12yz+8x+16y+24z+15$,
   (b) $x^2-6xy+10y^2+2xz-20z^2-10yz-40z-17$,
   (c) $x^2+4z^2+4xz+4x+4z-6y+6$,
   (d) $-2x^2-3y^2-7z^2+2xy-8yz-6xz-4x-6y-14z-6$.
3. 证明: 在二次函数 $\boldsymbol{XAX}^{\mathrm{T}}+\boldsymbol{BX}^{\mathrm{T}}+c$(其中 $\boldsymbol{A}$ 是非奇异矩阵) 中的线性部分可以通过平移消去.
4. (a) 证明: 非平凡实二次方程 $\boldsymbol{XAX}^{\mathrm{T}}=1$ 是一个旋转曲面当且仅当 $\boldsymbol{A}$ 具有二重特征值.
   (b) 描述二次方程 $xy+yz+zx=3$.
5. 把定理 23 中给出的二次函数的仿射分类推广到系数在任意满足 $1+1\neq 0$ 的域中的函数上.
6. (a) 列出三维空间中二次曲面的各种可能的仿射类型.
   (b) 对每一类型给出简短的几何描述.
7. 在广义相似群 (9.4 节末尾) 之下, 按 (a) 椭圆, (b) 抛物线, (c) 双曲线进行分类. 对每种情形求出全系不变量.

*8. 在刚体运动群之下, 对 n 维欧几里得空间中的二次超曲面进行分类 (利用定理 22).

9. 求在椭圆 $x^2+3y^2=3$ 中面积最大的内接六边形.

*9.12 酉矩阵与埃尔米特矩阵

对于复数的情况, 实二次型的正交变换是用某个埃尔米特型的酉变换来代替. 一个复数 $c = a + ib$ 定义为实数对 (a, b), 或定义为二维向量空间 $\mathbf{R}^2$ 中具有分量 (a, b) 的向量. 复数的模或绝对值 $|c|$ 正好是实向量的长度

$$|c|^2 = |a + ib|^2 = a^2 + b^2 = (a + ib)(a - ib) = cc^*, \tag{47}$$

这里 c^* 表示 c 的复共轭 $a - ib$. 同样道理, 具有 n 个复分量 $(c_1, \cdots, c_n)$(其中每个分量 $c_j = a_j + ib_j$) 的向量 γ 可以看作 $2n$ 维实空间中的具有 $2n$ 个分量 $(a_1, b_1, \cdots, a_n, b_n)$ 的向量. 这个实向量的长度由下式的平方根给出

$$\begin{aligned}|(c_1, \cdots, c_n)|^2 &= (a_1^2 + b_1^2) + \cdots + (a_n^2 + b_n^2)\\ &= \sum_{j=1}^{n}(a_j + ib_j)(a_j - ib_j)\\ &= c_1c_1^* + \cdots + c_nc_n^*.\end{aligned} \tag{48}$$

因为每个乘积 $c_jc_j^* = a_j^2 + b_j^2 \geqslant 0$, 所以这个表达式具有严格的正性: 这个实数和 $\sum_{j=1}^{n} c_jc_j^*$ 是正的, 除非所有的 $c_j = 0$. 表达式 (48) 很像通常的关于实向量长度的毕达哥拉斯公式. 我们用 (48) 式作为复行向量 $\boldsymbol{K} = (c_1 \cdots, c_n)$ 长度的定义. 公式 $\Sigma c_jc_j^*$ 可以写成矩阵形式 $\boldsymbol{K}\boldsymbol{K}^{*\mathrm{T}}$, 其中 $\boldsymbol{K}^*$ 是由向量 $\boldsymbol{K}$ 的每个分量取共轭而得到的向量.

定义 在复向量空间 $\mathbf{C}^n$ 中, 设 $\boldsymbol{\xi}$ 和 $\boldsymbol{\eta}$ 是具有坐标 $\boldsymbol{X} = (x_1, \cdots, x_n)$ 和 $\boldsymbol{Y} = (y_1, \cdots, y_n)$ 的向量, 并引进内积

$$(\boldsymbol{\xi}, \boldsymbol{\eta}) = x_1y_1^* + \cdots + x_ny_n^* = \boldsymbol{X}\boldsymbol{Y}^{*\mathrm{T}}, \tag{49}$$

那么, $\boldsymbol{\xi}$ 的长度是 $|\boldsymbol{\xi}| = (\boldsymbol{\xi}, \boldsymbol{\xi})^{\frac{1}{2}}$.

同普通内积的情况几乎一样, 我们可以证明这个内积具有基本性质:

线　　性　$(c\boldsymbol{\xi} + d\boldsymbol{\eta}, \boldsymbol{\zeta}) = c(\boldsymbol{\xi}, \boldsymbol{\zeta}) + d(\boldsymbol{\eta}, \boldsymbol{\zeta})$;

斜对称性　$(\boldsymbol{\xi}, \boldsymbol{\eta}) = (\boldsymbol{\eta}, \boldsymbol{\xi})^*$;

正　　性　如果 $\boldsymbol{\xi} \neq \mathbf{0}$, 则 $(\boldsymbol{\xi}, \boldsymbol{\xi})$是实数, 并且$(\boldsymbol{\xi}, \boldsymbol{\xi}) > 0$.

由斜对称性显然可推出第二个因子的斜线性:

$$\begin{aligned}(\boldsymbol{\xi}, c\boldsymbol{\eta} + d\boldsymbol{\zeta}) &= (c\boldsymbol{\eta} + d\boldsymbol{\zeta}, \boldsymbol{\xi})^* = c^*(\boldsymbol{\eta}, \boldsymbol{\xi})^* + d^*(\boldsymbol{\zeta}, \boldsymbol{\xi})^*\\ &= c^*(\boldsymbol{\xi}, \boldsymbol{\eta}) + d^*(\boldsymbol{\xi}, \boldsymbol{\zeta}).\end{aligned}$$

所以

$$(\boldsymbol{\xi}, c\boldsymbol{\eta}+d\boldsymbol{\zeta})=c^*(\boldsymbol{\xi},\boldsymbol{\eta})+d^*(\boldsymbol{\xi},\boldsymbol{\zeta}). \tag{50}$$

如果需要的话, 我们可以把线性、斜对称性和正性这些性质作为复数域上抽象向量空间中内积 $(\boldsymbol{\xi},\boldsymbol{\eta})$ 的公设. 那么这个空间被称为*酉空间*(比较 7.10 节的欧几里得向量空间).

两个向量 $\boldsymbol{\xi}$ 和 $\boldsymbol{\eta}$ 如果满足 $(\boldsymbol{\xi},\boldsymbol{\eta})=0$, 则称它们是*正交的*(记作 $\boldsymbol{\xi}\perp\boldsymbol{\eta}$). 根据斜对称性, $\boldsymbol{\xi}\perp\boldsymbol{\eta}$ 可推出 $\boldsymbol{\eta}\perp\boldsymbol{\xi}$. n 维复向量空间中一组 (n 个) 向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$, 如果每个向量的长度为 1, 且任意两个向量是正交的:

$$|\boldsymbol{\alpha}_1|=\cdots=|\boldsymbol{\alpha}_n|=1,\quad (\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)=0\quad (i\neq j), \tag{51}$$

则称这 n 个向量是这个空间的*标准酉基*. 这样一组向量一定是普通意义下的基. 原来的基向量 $\boldsymbol{\epsilon}_1=(1,0,\cdots,0),\cdots,\boldsymbol{\epsilon}_n=(0,\cdots,0,1)$ 构成一组标准酉基. 根据 7.11 节的方法, 我们还可以构造另外的标准酉基, 并可证明

定理 24 *在酉空间中, 任意 $m(<n)$ 个长度为 1 的相互正交的向量构成这个空间标准酉基的一部分.*

特别是, 如果 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 是非零正交向量, $c_i=\dfrac{(\boldsymbol{\xi},\alpha_i)}{(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_i)}$, 那么对任意 $\boldsymbol{\xi}$, $\boldsymbol{\alpha}_{m+1}=\boldsymbol{\xi}-c_1\boldsymbol{\alpha}_1-\cdots-c_m\boldsymbol{\alpha}_m$ 都同向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 正交.

一个 $n\times n$ 复矩阵 $\boldsymbol{U}=(u_{ij})$, 如果满足 $\boldsymbol{U}\boldsymbol{U}^{*\mathrm{T}}=\boldsymbol{I}$(其中 $\boldsymbol{U}^*$ 表示由 $\boldsymbol{U}$ 的每个元素取共轭而得到的矩阵), 则称 $\boldsymbol{U}$ 为*酉矩阵*. 这个条件显然等价于 $\sum\limits_k u_{ik}u_{jk}^*=\delta_{ij}$, 这里 δ_{ij} 是克罗内克尔符号 (9.4 节). 换句话说, $\boldsymbol{U}$ 的每一行向量的长度为 1, $\boldsymbol{U}$ 的任意两个行向量正交. 这意味着, 由 $\boldsymbol{U}$ 定义的 $\mathbf{C}^n$ 的线性变换把 $\boldsymbol{\epsilon}_1,\cdots,\boldsymbol{\epsilon}_n$ 变换到一组标准酉基. 根据 8.6 节定理 9 的推论 6, 它还等价于条件 $\boldsymbol{U}^{*\mathrm{T}}\boldsymbol{U}=\boldsymbol{I}$, 这表明 $\boldsymbol{U}$ 的每一列向量的长度为 1, $\boldsymbol{U}$ 的任意两个列向量正交.

$\mathbf{C}^n$ 的任意线性变换 $\boldsymbol{X}\longmapsto\boldsymbol{X}\boldsymbol{A}$ 把内积 $\boldsymbol{X}\boldsymbol{Y}^{*\mathrm{T}}$ 变换成 $\boldsymbol{X}\boldsymbol{A}\boldsymbol{A}^{*\mathrm{T}}\boldsymbol{Y}^{*\mathrm{T}}$. 对所有的向量 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$, 这个新的内积等于原来的内积 $\boldsymbol{X}\boldsymbol{Y}^{*\mathrm{T}}=\boldsymbol{X}\boldsymbol{I}\boldsymbol{Y}^{*\mathrm{T}}$ 当且仅当 $\boldsymbol{A}\boldsymbol{A}^{*\mathrm{T}}=\boldsymbol{I}$, 即当且仅当 $\boldsymbol{A}$ 是酉矩阵. 于是, 一个矩阵 $\boldsymbol{A}$ 是酉矩阵当且仅当 $\boldsymbol{A}$ 所对应的线性变换 $T_{\boldsymbol{A}}$ 保持复内积 $\boldsymbol{X}\boldsymbol{Y}^{*\mathrm{T}}$ 不变. 类似的论证指出, $\boldsymbol{A}$ 是酉矩阵当且仅当 $T_{\boldsymbol{A}}$ 保持长度 $(\boldsymbol{X}\boldsymbol{X}^{*\mathrm{T}})^{\frac{1}{2}}$ 不变. 几何上, 如果酉空间的线性变换 T 保持长度不变, $|\boldsymbol{\xi}T|=|\boldsymbol{\xi}|$(因而也保持内积不变), 那么称 T 为*酉变换*. n 维空间的所有酉变换组成的集合构成一个群, 这个群与所有 $n\times n$ 酉矩阵组成的群同构.

下面, 我们用 "埃尔米特型" 来代替二次型, 它的最简单的例子是长度公式 $\Sigma x_ix_i^*$. 一般地, *埃尔米特型*是一个含有复系数 h_{ij} 的表达式

$$\sum_{i,j=1}^n x_ih_{ij}x_j^*=\boldsymbol{X}\boldsymbol{H}\boldsymbol{X}^{*\mathrm{T}},\quad \boldsymbol{H}=(h_{ij}), \tag{52}$$

其中系数矩阵 $\boldsymbol{H}$ 具有性质 $\boldsymbol{H}^{*\mathrm{T}}=\boldsymbol{H}$. 这种类型的矩阵 $\boldsymbol{H}$ 称为埃尔米特矩阵. 当元素 h_{ij} 都是实数时, 埃尔米特矩阵就是对称矩阵. 埃尔米特型 (52) 可以看作对于某基的坐标为 $x_1,\cdots,x_n$ 的向量 $\boldsymbol{\xi}$ 的函数 $h(\boldsymbol{\xi})=\boldsymbol{XHX}^{*\mathrm{T}}$. 这个函数的值 $\boldsymbol{XHX}^{*\mathrm{T}}$ 总是实数. 为证明这一结论, 只须证明这个数值等于它的共轭 (或者证明等于它的共轭转置). 而因为 $\boldsymbol{H}$ 是埃尔米特矩阵, 所以有

$$(\boldsymbol{XHX}^{*\mathrm{T}})^{*\mathrm{T}}=(\boldsymbol{X}^*\boldsymbol{H}^*\boldsymbol{X}^{**\mathrm{T}})^{\mathrm{T}}=(\boldsymbol{X}^{\mathrm{T}})^{\mathrm{T}}\boldsymbol{H}^{*\mathrm{T}}\boldsymbol{X}^{*\mathrm{T}}=\boldsymbol{XHX}^{*\mathrm{T}},$$

满足断言.

酉变换 $\boldsymbol{Y}=\boldsymbol{XU},\ \boldsymbol{X}=\boldsymbol{YU}^{-1}=\boldsymbol{YU}^{*\mathrm{T}}$ 应用到埃尔米特型, 得到

$$\boldsymbol{XHX}^{*\mathrm{T}}=(\boldsymbol{YU}^{-1})\boldsymbol{H}(\boldsymbol{YU}^{*\mathrm{T}})^{*\mathrm{T}}=\boldsymbol{YU}^{-1}\boldsymbol{H}(\boldsymbol{UY}^{*\mathrm{T}})=\boldsymbol{Y}(\boldsymbol{U}^{-1}\boldsymbol{HU})\boldsymbol{Y}^{*\mathrm{T}}.$$

这个系数矩阵 $\boldsymbol{U}^{-1}\boldsymbol{HU}$ 还是埃尔米特矩阵, 这因为由于 $\boldsymbol{U}^{-1}=\boldsymbol{U}^{*\mathrm{T}}$, 有

$$(\boldsymbol{U}^{-1}\boldsymbol{HU})^{*\mathrm{T}}=\boldsymbol{U}^{*\mathrm{T}}\boldsymbol{H}^{*\mathrm{T}}(\boldsymbol{U}^{-1})^{*\mathrm{T}}=\boldsymbol{U}^{-1}\boldsymbol{HU}.$$

如果我们用坐标变换, 变换到一个新的标准酉坐标系, 那么可以得到同样的效果, 因为这样的变换通过方程 $\boldsymbol{Y}=\boldsymbol{XU}$(其中 $\boldsymbol{U}$ 是酉矩阵) 给出 $\boldsymbol{\xi}$ 的新坐标 $\boldsymbol{Y}$ 与老坐标的关系.

用这种变换的说法, 我们可以把任意埃尔米特型变换到主轴上. 这个新主轴是根据逐次最大的性质而选取的, 这恰好同二次型在正交变换下化到主轴的讨论一样. 第一轴 $\boldsymbol{\alpha}_1$ 取作长度为 1 的向量, 并使 $h(\boldsymbol{\xi})$ 在 $|\boldsymbol{\xi}|=1$ 上取最大值; 然后我们根据定理 24 可以求出包含 $\boldsymbol{\alpha}_1$ 在内的标准酉基. 对于这组基, 交叉乘积项 $x_1x_j^*(j\neq 1)$ 又被去掉. 因为埃尔米特型的值总是实的, 所以逐次最大值 λ_i 都是实数. 这一过程证明了下面的主轴定理.

定理 25 任意埃尔米特型 $\boldsymbol{XHX}^{*\mathrm{T}}$ 可以通过酉变换 $\boldsymbol{Y}=\boldsymbol{XU}$ 化成实对角型

$$\boldsymbol{YHY}^{*\mathrm{T}}=\lambda_1y_1y_1^*+\lambda_2y_2y_2^*+\cdots+\lambda_ny_ny_n^*. \tag{53}$$

这个定理可以平移到埃尔米特型的矩阵 $\boldsymbol{H}$ 上, 即

定理 26 对于每个埃尔米特矩阵 $\boldsymbol{H}$, 存在一个酉矩阵 $\boldsymbol{U}$, 使得 $\boldsymbol{U}^{-1}\boldsymbol{HU}=\boldsymbol{U}^{*\mathrm{T}}\boldsymbol{HU}$ 是实对角矩阵.

用第 10 章的方法还可证明 (53) 式的系数 λ_i 是唯一的.

习 题

1. 下列矩阵中哪些是酉矩阵? 哪些是埃尔米特矩阵?

$$\begin{pmatrix} \frac{1+\mathrm{i}}{2} & \frac{1-\mathrm{i}}{2} \\ \frac{1-\mathrm{i}}{2} & \frac{1+\mathrm{i}}{2} \end{pmatrix}, \quad \begin{pmatrix} 3 & 1-\mathrm{i} \\ 1+\mathrm{i} & \sqrt{2} \end{pmatrix}, \quad \begin{pmatrix} 1 & \mathrm{i} \\ \mathrm{i} & 1 \end{pmatrix}.$$

2. 求与 $\left(\frac{1}{2}, \frac{\mathrm{i}}{2}, \frac{1+\mathrm{i}}{2}\right)$ 正交的向量组成的子空间的一组标准酉基.
3. 证明：(h_{ij}) 是埃尔米特矩阵当且仅当对所有的 i, j 有 $h_{ij}^* = h_{ji}$.
4. 证明：如果 ω 是 n 次本原单位根，那么 $n^{-\frac{1}{2}}(\omega^{ij})(i, j = 1, \cdots, n)$ 是酉矩阵.
5. 证明：对任意实数 θ, 复矩阵

$$\begin{pmatrix} \mathrm{ch}\,\theta & \mathrm{i}\,\mathrm{sh}\,\theta \\ -\mathrm{i}\,\mathrm{sh}\,\theta & \mathrm{ch}\,\theta \end{pmatrix}$$

是酉矩阵. 计算它的特征值和特征向量.
6. 证明：所有 $n \times n$ 酉矩阵构成一个群 (酉群), 这个群与所有 $2n \times 2n$ 实正交矩阵组成的群的一个子群同构.
7. 证明：埃尔米特内积 $(\boldsymbol{\xi}, \boldsymbol{\eta})$ 满足线性、斜对称性和正性.
8. 详细证明关于标准酉基的定理 24.
9. 证明：一个单项矩阵是酉矩阵当且仅当它的所有非零元素的绝对值是 1.
10. 对于两个变量的埃尔米特型，它在 $x = 0, y = 1$ 点处取最大值，证明类似于 9.10 节引理的结果. (提示：把每个变量分成实部和虚部.)

*11. 给出埃尔米特型主轴定理的详细证明.

12. 通过关于 x 和 y 的酉变换，把埃尔米特型 $xy^* + x^*y$ 化成对角型. (提示：考虑相应的实二次型.)
13. 在酉群之下，把 $zz^* - 2ww^* + 2\mathrm{i}(zw^* - wz^*)$ 化成对角型.
14. 证明：任意实斜对称矩阵 $\boldsymbol{A}$ 具有一组由复特征向量组成的基，这些特征向量对应的特征值都是纯虚数. (提示：证明 $\mathrm{i}\,\boldsymbol{A}$ 是埃尔米特矩阵.)
15. 证明：任意酉矩阵的谱位于复平面的单位圆上.
16. 证明：复矩阵 $\boldsymbol{C}$ 是正定的埃尔米特矩阵当且仅当 $\boldsymbol{C} = \boldsymbol{P}\boldsymbol{P}^{*\mathrm{T}}$, 其中 $\boldsymbol{P}$ 是某非奇异矩阵.
17. 证明：埃尔米特矩阵是正定的当且仅当它的所有特征值都是正的.

*9.13 仿射几何

仿射几何是研究在仿射群之下图形不变的性质. 正如欧几里得几何是研究在欧几里得群之下图形不变的性质一样. 作用在有限维向量空间 V 上的仿射群，像 (11) 式那样，是由 V 的所有把 V 的点 (或向量)$\boldsymbol{\xi}$ 变到点

$$\boldsymbol{\xi}H = \boldsymbol{\eta} = \boldsymbol{\xi}T + \boldsymbol{\kappa} \tag{54}$$

的变换 H 组成的, 这里 $\boldsymbol{\kappa}$ 是一固定向量, T 是 V 的一固定非奇异线性变换. 我们假定 V 是域 F 上的向量空间, 其中域 F 满足条件 $1+1\neq 0$(例如 F 不是域 $\mathbf{Z}_2$).

在仿射几何中, 正如在欧几里得几何中一样, 任意两个点 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 是等价的, 这因为平移 $\boldsymbol{\xi}\longmapsto\boldsymbol{\xi}+(\boldsymbol{\beta}-\boldsymbol{\alpha})$ 把 $\boldsymbol{\alpha}$ 变到 $\boldsymbol{\beta}$. 这就把仿射几何与 V 的向量几何 (在全线性群之下) 区别开来, 在仿射几何中原点 $\boldsymbol{O}$ 起着 V 中零向量 $\mathbf{0}$ 的特殊作用. 当考虑在仿射群之下保持不变的性质时, 我们通常把向量空间称作仿射空间.

在平面解析几何中, 联结两点 (x_1,y_1) 和 (x_2,y_2) 的直线的方程是

$$y-y_1=\frac{y_2-y_1}{x_2-x_1}(x-x_1),\quad x_2\neq x_1.$$

引入参数 $t=\dfrac{x-x_1}{x_2-x_1}$, 则我们得到 $y=y_1+t(y_2-y_1)$, $x=x_1+t(x_2-x_1)$; 换句话说, 这直线具有参数方程

$$x=(1-t)x_1+tx_2,\qquad y=(1-t)y_1+ty_2, \tag{55}$$

它可以写成向量形式 $(x,y)=\boldsymbol{\xi}=(1-t)\boldsymbol{\xi}_1+t\boldsymbol{\xi}_2$. 几何上, (55) 的点 (x,y) 是把 (x_1,y_1) 和 (x_2,y_2) 两点之间的线段分成比例为 $t:(1-t)$ 的点, 当 $t=\dfrac{1}{2}$, 这个点就是中点.

在任意仿射空间中, 把 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 之间的 "线段" 分成比例为 $t:(1-t)$ 的点定义为

$$\boldsymbol{\gamma}=(1-t)\boldsymbol{\alpha}+t\boldsymbol{\beta}. \tag{56}$$

当 $\boldsymbol{\alpha}\neq\boldsymbol{\beta}$, 联结 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 的 (仿射) 直线 $\overline{\boldsymbol{\alpha\beta}}$ 定义为满足 (56) 式 $(t\in F)$ 的所有点 $\boldsymbol{\gamma}$ 组成的集合.

定理 27 任意非奇异仿射变换把直线变到直线.

证明 把 (56) 代入 (54) 中, 我们有

$$\begin{aligned}\boldsymbol{\gamma}H&=\boldsymbol{\gamma}T+\boldsymbol{\kappa}=(1-t)\boldsymbol{\alpha}T+t\boldsymbol{\beta}T+\boldsymbol{\kappa}\\&=(1-t)(\boldsymbol{\alpha}T+\boldsymbol{\kappa})+t(\boldsymbol{\beta}T+\boldsymbol{\kappa})=(1-t)(\boldsymbol{\alpha}H)+t(\boldsymbol{\beta}H).\end{aligned}$$

因此 H 把通过 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 的仿射直线 $\overline{\boldsymbol{\alpha\beta}}$ 变到通过 $\boldsymbol{\alpha}H$ 和 $\boldsymbol{\beta}H$ 的仿射直线. 证毕

如果 $\boldsymbol{\gamma}=(1-t)\boldsymbol{\alpha}+t\boldsymbol{\beta}$ 和 $\boldsymbol{\delta}=(1-u)\boldsymbol{\alpha}+u\boldsymbol{\beta}$ 是 $\overline{\boldsymbol{\alpha\beta}}$ 上任意两个不同的点, 那么因为

$$(1-v)\boldsymbol{\gamma}+v\boldsymbol{\delta}=(1-t+vt-vu)\boldsymbol{\alpha}+(t-vt+vu)\boldsymbol{\beta},$$

所以 $\overline{\boldsymbol{\alpha\beta}}$ 包含 $\overline{\boldsymbol{\gamma\delta}}$ 的每个点. 反过来也可以类似地证明 $\overline{\boldsymbol{\gamma\delta}}$ 包含 $\overline{\boldsymbol{\alpha\beta}}$ 的每个点, 因此 $\overline{\boldsymbol{\alpha\beta}}=\overline{\boldsymbol{\gamma\delta}}$. 也就是说, 直线可用它上面的任意两点确定.

一个普通平面有时可用平直性质来表征：平面如果包含任意两点，则它必包含通过这两点的整个直线. 我们可以用这种性质把 V 的仿射子空间定义为 V 中具有下面性质的任意子集合 M：如果 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 在 M 中，则整个直线 $\overline{\boldsymbol{\alpha\beta}}$ 也在 M 中. 显然，仿射变换把仿射子空间映上到仿射子空间. 更进一步，V 的仿射子空间恰恰是按下述意义平移 V 的向量子空间而得到的一个子空间.

定理 28 *如果 M 是 V 的任意仿射子空间，那么存在 V 的一个线性子空间 S 和向量 $\boldsymbol{\kappa}$，使得 M 是由所有点 $\boldsymbol{\xi}+\boldsymbol{\kappa}$ 组成，其中 $\boldsymbol{\xi}\in S$. 反过来，任意 S 和 $\boldsymbol{\kappa}$ 可以按这种方法确定一个仿射子空间 $M=S+\boldsymbol{\kappa}$.*

证明 设 $\boldsymbol{\kappa}$ 是 M 中任意一点，并定义 S 是所有向量 $\boldsymbol{\alpha}-\boldsymbol{\kappa}$ 组成的集合 (其中 $\boldsymbol{\alpha}\in M$)；换句话说，S 是按 $-\boldsymbol{\kappa}$ 平移 M 而得到的. 显然，M 具有所要求的按 S 和 $\boldsymbol{\kappa}$ 表达的形式. 剩下只须证明 S 是一个向量子空间. 因为直线平移到直线，所以对于 M 的假设保证 S 也有同样的性质：联结 S 的任意两个向量的直线仍在 S 中. 对于 S 中任意向量 $\boldsymbol{\alpha}$，联结 $\boldsymbol{O}(\in S)$ 和 $\boldsymbol{\alpha}$ 的直线在 S 中，因此 S 包含所有 “数乘” 积 $c\boldsymbol{\alpha}$. 如果 S 包含 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$，那么它必包含 $2\boldsymbol{\alpha}$ 和 $2\boldsymbol{\beta}$ 以及联结它们的整个直线 $\boldsymbol{\xi}=2\boldsymbol{\alpha}+t(2\boldsymbol{\beta}-2\boldsymbol{\alpha})$. (画一个图!) 特别当 $t=\dfrac{1}{2}$ 时，S 包含 $\boldsymbol{\xi}=2\boldsymbol{\alpha}+(\boldsymbol{\beta}-\boldsymbol{\alpha})=\boldsymbol{\beta}+\boldsymbol{\alpha}$，也就是包含给定向量的和. 于是，我们就证明了 S 在加法和数乘运算之下是封闭的，因此它是向量子空间，满足要求. 证毕

$F=\mathbf{Z}_2$ 的情形是个例外：三向量组 $(0,0),(1,0)$ 和 $(0,1)$ 是一个 “平面”，当它包含任意两点 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$，就必包含所有 $(1-t)\boldsymbol{\alpha}+t\boldsymbol{\beta}$；但是这个三向量组并不是一个仿射子空间.

逆定理也不难证明. 它换种说法就是，仿射子空间正好是在向量加法群之下向量子空间的陪集. 特别是，仿射直线是一维向量子空间的陪集 (在平移之下).

上述结果包含着仿射几何中另一个概念：平行性.

定义 *仿射空间 V 的两个子集合 S 和 S^* 称为是平行的当且仅当存在 V 的一个平移 $L:\boldsymbol{\xi}\longmapsto\boldsymbol{\xi}+\boldsymbol{\lambda}$ 把 S 映上到 S^*.*

定理 29 *V 的任意仿射变换把平行集合映射到平行集合.*

证明 设 S 和 $S^*=S+\boldsymbol{\lambda}$ 是给定的平行集合，设 U 和 U^* 分别是它们在 $H:\boldsymbol{\xi}\longmapsto\boldsymbol{\xi}T+\boldsymbol{\kappa}$ 之下的变换式. 定理断言，U^* 是由所有 $\boldsymbol{\xi}+\boldsymbol{\mu}$ 组成的集合，其中 $\boldsymbol{\xi}\in U$，$\boldsymbol{\mu}$ 是某一固定的平移向量. 根据定义，U^* 是由所有 $(\boldsymbol{\sigma}+\boldsymbol{\lambda})T+\boldsymbol{\kappa}=(\boldsymbol{\sigma}T+\boldsymbol{\kappa})+\boldsymbol{\lambda}T$ 组成的集合，其中 $\sigma\in S$. 并且 U 是由所有 $\boldsymbol{\xi}=\boldsymbol{\sigma}T+\boldsymbol{\kappa}$ 组成的集合，其中 $\boldsymbol{\sigma}\in S$. 置 $\boldsymbol{\mu}=\boldsymbol{\lambda}T$，显然定理的结论成立. 证毕

在仿射群 (实数域 $\mathbf{R}$ 上) 之下的等价性在初等几何上有很多有趣的应用. 在仿射群之下，任意两个三角形是等价的. 为了证明这一点，只须证明任意三角形 $\boldsymbol{\alpha\beta\gamma}$ 等价于以 $\boldsymbol{O}=(0,0),\boldsymbol{\beta}_0=(2,0),\boldsymbol{\gamma}_0=(1,\sqrt{3})$ 为顶点的特殊的等边三角形 (见

图 9-2). 通过平移, 顶点 $\boldsymbol{\alpha}$ 可以移动到原点 $\boldsymbol{O}$, 其他两顶点移到位置 $\boldsymbol{\beta}'$ 和 γ'. 因为向量 $\boldsymbol{\beta}'$ 和 γ' 是线性无关的, 则存在一个线性变换 $x\boldsymbol{\beta}' + y\gamma' \longmapsto x\boldsymbol{\beta}_0 + y\gamma_0$, 它把 $\boldsymbol{\beta}'$ 变到 $\boldsymbol{\beta}_0$, 把 γ' 变到 γ_0. 这个变换与平移的乘积将把 $\boldsymbol{\alpha\beta}\gamma$ 变到 $\boldsymbol{O\beta}_0\gamma_0$, 正如所求. 因此这两个三角形是等价的.

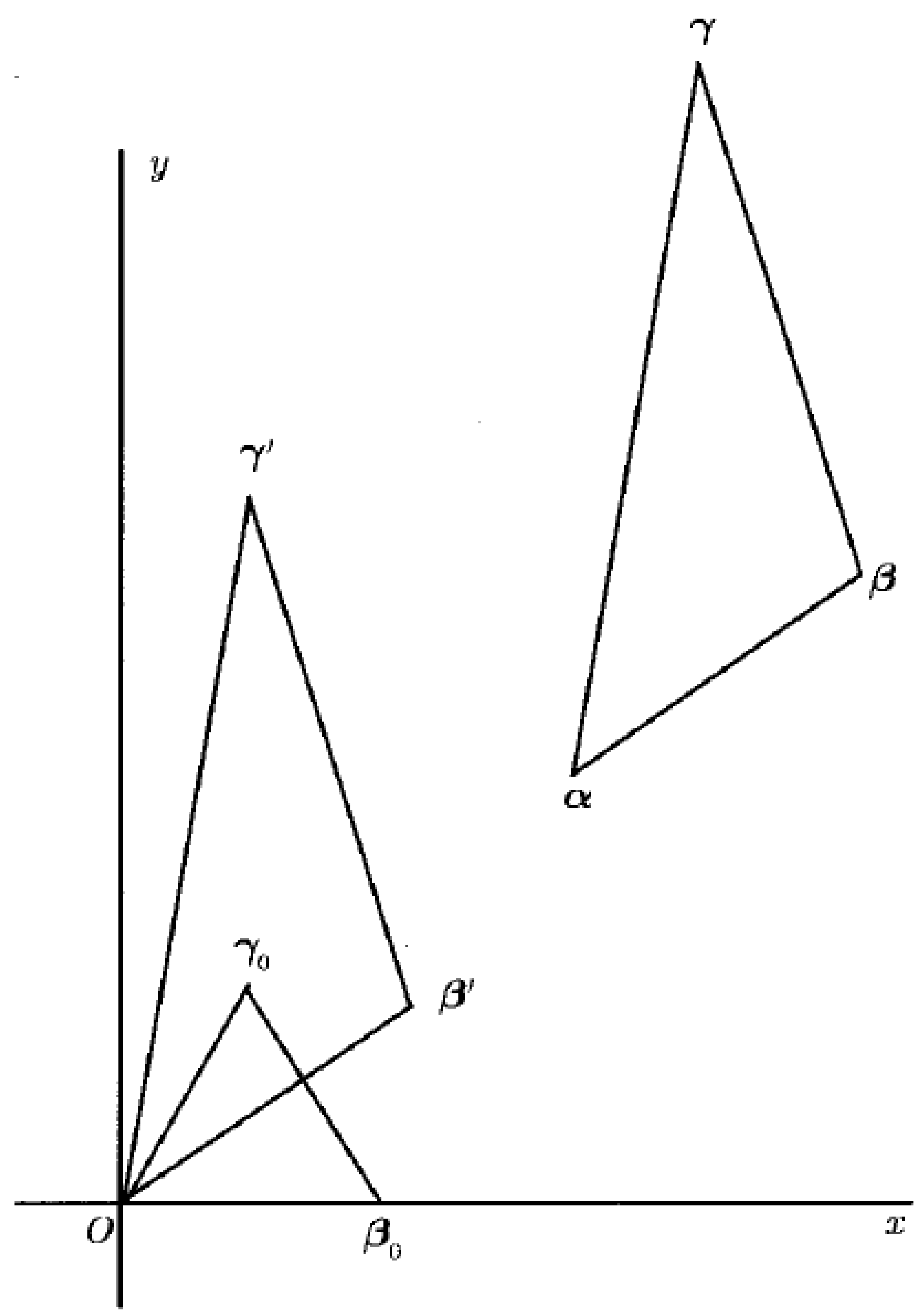


图 9-2

于是每个三角形等价于一个等边三角形. 而根据对称性, 等边三角形的三条中线一定交于一点 (即重心). 但是, 仿射变换把中点变到中点, 因此它把中线变到中线. 这就证明了一个初等定理: 任意三角形三中线交于一点. 再有, 我们可以很容易地证明, 等边三角形中线的交点按比例 1:2 把中线分成两部分, 因此对任意三角形同样的结论成立.

此外, 任意椭圆与一个圆仿射等价. 但是一个圆的通过圆心的任意直径在两个端点上有互相平行的切线, 而且平行于这两条切线的共轭直径把平行于给定直径的所有弦平分. 可以推出, 对任意椭圆, 这两个性质同样成立, 这是因为仿射变换把平行线变为平行线, 把切线变为切线 (但应注意, 椭圆中的共轭直径不一定互相垂直, 图 9-3).

附录 形心与重心坐标 按给定的比例分割一线段的分点 (56) 是形心概念的特殊情形. 在 V 中给定 $m+1$ 个点 $\boldsymbol{\alpha}_0, \cdots, \boldsymbol{\alpha}_m$, 在 F 中给定 $m+1$ 个元素 $x_0, \cdots, x_m$,

使得 $x_0+\cdots+x_m=1$, 具有权 $x_0,\cdots,x_m$ 的点 $\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_m$ 的形心定义为点

$$\boldsymbol{\xi}=x_0\boldsymbol{\alpha}_0+\cdots+x_m\boldsymbol{\alpha}_m,\ x_0+\cdots+x_m=1. \tag{57}$$

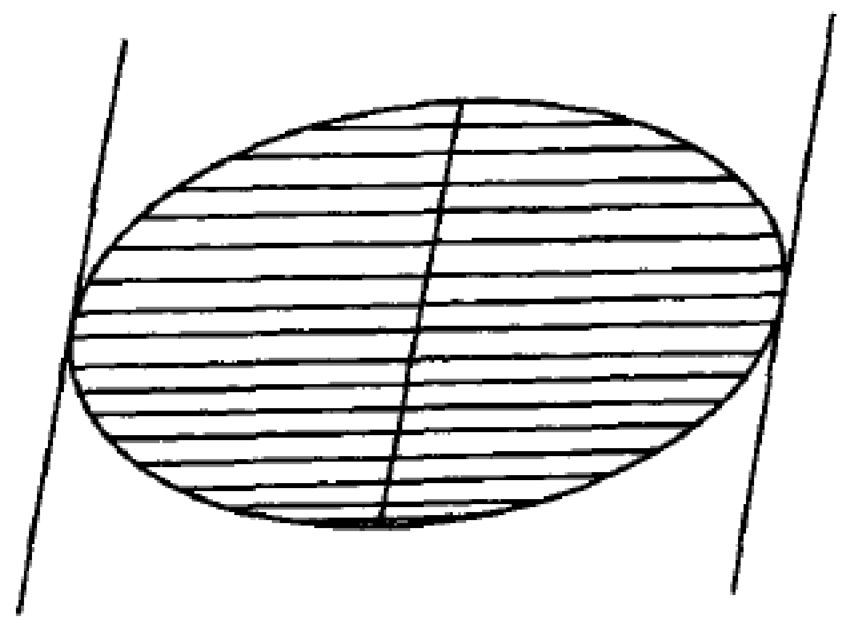


图 9-3

$\Big($更一般地, 对任意 $w=w_0+\cdots+w_m\neq 0$, 具有权 $w_0,\cdots,w_m$ 的点 $\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_m$ 的形心仍用 (57) 来定义, 这时 $x_i=\dfrac{w_i}{w}$.$\Big)$

如果 H 是形为 (54) 的任意仿射变换, 那么

$$\begin{aligned}\boldsymbol{\xi}H&=(x_0\boldsymbol{\alpha}_0+\cdots+x_m\boldsymbol{\alpha}_m)T+\boldsymbol{\kappa}\\&=x_0(\boldsymbol{\alpha}_0T)+\cdots+x_m(\boldsymbol{\alpha}_mT)+(\Sigma x_i)\boldsymbol{\kappa}\\&=x_0(\boldsymbol{\alpha}_0H)+\cdots+x_m(\boldsymbol{\alpha}_mH).\end{aligned}$$

换句话说, 仿射变换把形心变换到形心, 其权不变.

定理 30 仿射子空间 M 包含它的点的所有形心.

我们通过对 (57) 式的点数 $m+1$ 用归纳法来证明这个定理. 当 $m=0$, 可直接得出结论. 当 $m=1$, $\boldsymbol{\alpha}_0$ 和 $\boldsymbol{\alpha}_1$ 的形心恰好是通过 $\boldsymbol{\alpha}_0$ 和 $\boldsymbol{\alpha}_1$ 的直线上的一个点, 因此根据定义, 这个形心在 M 中. 假定 $m>1$, 并把 $\boldsymbol{\xi}$ 看作形为 (57) 的点. 那么有某个系数, 比如说是 x_m 不等于 1. 设 $t=x_0+\cdots+x_{m-1}$, 则 $x_m=1-t, t\neq 0$, 点 $\boldsymbol{\beta}=\dfrac{x_0}{t}\boldsymbol{\alpha}_0+\cdots+\dfrac{x_{m-1}}{t}\boldsymbol{\alpha}_{m-1}$ 是 $\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_{m-1}$ 的形心, 由归纳法假设, 它在 M 中. 进一步, $\boldsymbol{\xi}=t\boldsymbol{\beta}+(1-t)\boldsymbol{\alpha}_m$ 是在联结 $\boldsymbol{\beta}\in M$ 和 $\boldsymbol{\alpha}_m\in M$ 的直线上, 因此 $\boldsymbol{\xi}$ 在 M 中, 如断言所述.

形心可用来描述由给定的点集 $\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_m$ 张成的子空间 M, 如下所述.

定理 31 V 中 $m+1$ 个点 $\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_m$ 的所有形心 (57) 组成的集合是一个仿射子空间 M. 这个子空间 M 包含每个点 $\boldsymbol{\alpha}_i$, 并且它包含在任意仿射子空间 N 中, 而 N 包含所有的点 $\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_m$.

证明 设 (57) 表示的 $\boldsymbol{\xi}$ 和

$$\boldsymbol{\eta}=y_0\boldsymbol{\alpha}_0+\cdots+y_m\boldsymbol{\alpha}_m,\ y_0+\cdots+y_m=1 \tag{57'}$$

是任意两个形心, 则

$$(1-t)\boldsymbol{\xi}+t\boldsymbol{\eta}=[(1-t)x_0+ty_0]\boldsymbol{\alpha}_0+\cdots+[(1-t)x_m+ty_m]\boldsymbol{\alpha}_m$$

也是 $\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_m$ 的形心, 这是因为全体系数 $(1-t)x_i+ty_i$ 的和是 1. 因此 M 的确是一个仿射子空间. 显然它包含每个 $\boldsymbol{\alpha}_i$. 另一方面, 根据定理 30, 包含所有 $\boldsymbol{\alpha}_i$ 的任意仿射子空间 N, 一定包含整个 M. 证毕

如果 m 个向量 $\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_m-\boldsymbol{\alpha}_0$ 是线性无关的, 那么称 $m+1$ 个点 $\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_m$ 是仿射无关的. 对于一个仿射变换 H, 我们有 $(\boldsymbol{\alpha}_i-\boldsymbol{\alpha}_0)T=\boldsymbol{\alpha}_iH-\boldsymbol{\alpha}_0H$, 因此非奇异仿射变换把仿射无关的点变到仿射无关的点. 在这个仿射无关的定义中, 起点 $\boldsymbol{\alpha}_0$ 起着特殊的作用. 下面的结果表明, 仿射无关性不依赖于起点的选取.

定理 32　*$m+1$ 个点仿射无关当且仅当由 $\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_m$ 张成的仿射子空间 M 中的每一点 $\boldsymbol{\xi}$, 作为 $\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_m$ 的形心 (57), 有唯一的表达式.*

证明　假设点 $\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_m$ 仿射无关, 而 M 中某一点 $\boldsymbol{\xi}$, 作为形心有两种表达式 $\boldsymbol{\xi}=\Sigma x_i\boldsymbol{\alpha}_i$, $\boldsymbol{\xi}=\Sigma x_i'\boldsymbol{\alpha}_i$, 并满足 $\Sigma x_i=1=\Sigma x_i'$. 那么

$$x_0'-x_0=(x_1-x_1')+\cdots+(x_m-x_m'),$$

而且零向量 $\mathbf{0}=\boldsymbol{O}$ 具有表达式

$$\begin{aligned}\mathbf{0}&=\sum_{i=0}^{m}(x_i-x_i')\boldsymbol{\alpha}_i=\sum_{i=1}^{m}(x_i-x_i')\boldsymbol{\alpha}_i-(x_0'-x_0)\boldsymbol{\alpha}_0\\&=\sum_{i=1}^{m}(x_i-x_i')(\boldsymbol{\alpha}_i-\boldsymbol{\alpha}_0)\end{aligned}$$

因为向量 $\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_m-\boldsymbol{\alpha}_0$ 是线性无关的, 我们推出结论: 对于 $i=1,\cdots,m$ 有 $x_i=x_i'$. 因为 $x_0=1-(x_1+\cdots+x_m)$, 所以我们还有 $x_0=x_0'$. 于是, $\boldsymbol{\xi}$ 作为形心的表达式是唯一的.

其次, 假设点 $\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_m$ 仿射相关. 那么存在一个线性关系 $\Sigma c_i(\boldsymbol{\alpha}_i-\boldsymbol{\alpha}_0)=\mathbf{0}$, 其中系数 c_i 不全为零, 比如说 $c_1\neq 0$. 通过做除法, 我们可以假定 $c_1=1$. 则有

$$\boldsymbol{\alpha}_1=-c_2\boldsymbol{\alpha}_2-\cdots-c_m\boldsymbol{\alpha}_m+(c_2+\cdots+c_m+1)\boldsymbol{\alpha}_0,$$

在 $\boldsymbol{\alpha}_1$ 的表达式中, 这些系数之和是 1, 但是 $\boldsymbol{\alpha}_1$ 还有另一种表达式 $\boldsymbol{\alpha}_1=1\cdot\boldsymbol{\alpha}_1$, 因此 $\boldsymbol{\alpha}_1$ 作为形心的表达式不是唯一的. 证毕

当点 $\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_m$ 是仿射无关时, 由 $\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_m$ 张成的空间中的每一点可表成 (57), 在表达式 (57) 中出现的标量 $x_0,\cdots,x_m$ 称为 $\boldsymbol{\xi}$ 相对于 $\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_m$ 的**重心坐标**. 注意, 这些坐标中任意 m 个可以确定剩下的一个坐标, 这是由于 $x_0+\cdots+x_m=1$.

习 题

1. 对下列每对点, 求两点联线的参数方程, 并把这直线表示成形式 $S_1+\lambda$(也就是求空间 S_1).

   (a) $(2,1)$ 和 $(5,0)$, (b) $(1,3,2)$ 和 $(-1,7,5)$, (c) $(1,2,3,4)$ 和 $(4,3,2,1)$.

2. 把通过 $(1,3)$ 和 $(4,2)$ 两点的直线表示成形式 $S+\lambda$, 其中 λ 有四种不同的选择. 画出图形.

3. 证明: 通过不在一直线上的三个向量 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$, 存在一个且只存在一个二维仿射子空间 (一个平面!). 证明: 这个平面上的向量可表示成形式 $\boldsymbol{\xi}=\boldsymbol{\alpha}+s(\boldsymbol{\beta}-\boldsymbol{\alpha})+t(\boldsymbol{\gamma}-\boldsymbol{\alpha})$, 其中 s 和 t 是变量.

4. 求通过下列每三点的平面 (如果存在的话) 的参数方程 (按照习题 3 的形式):

   (a) $(1,3,2)$, $(4,1,-1)$, $(2,0,0)$,

   (b) $(1,1,0)$, $(1,0,1)$, $(0,1,1)$,

   (c) $(2,-1,3)$, $(1,1,1)$, $(3,0,4)$.

5. 在习题 4 的每个小题中, 求通过原点的平行平面的基.

*6. 证明: 定理 28 在除去 $\mathbf{Z}_2$ 的每个域上都成立.

7. 只用有关的定义证明: 任意仿射变换把中点变到中点.

8. 证明: 每个平行四边形与正方形仿射等价.

9. 用仿射方法证明: 平行四边形的两条对角线总是彼此平分.

10. (a) 求 $\mathbf{R}^2$ 的仿射变换, 它把以 (0,0), (0,1), (1,0) 为顶点的三角形变换到以 $(1,0)$, $(-1,0)$, $(0,\sqrt{3})$ 为顶点的等边三角形.

    (b) 如果第一个三角形是以 (1,1), (1,2), (3,3) 为顶点, 那么所求的仿射变换是什么?

11. 用仿射方法证明: 在梯形中, 两条对角线和两平行边中点联线通过一点.

12. 证明: 任意平行六面体与正立方体仿射等价.

13. 证明: 任意平行六面体的四条对角线有公共的中点 (它是重心).

14. (a) 证明: 在任意域 F 上, 任意两个三角形在仿射群之下是等价的.

    *(b) 证明: 如果域 F 中 $1+1\neq 0, 1+1+1\neq 0$, 那么任意三角形的中线交于一点.

15. 证明: 向量空间 V 的一一变换 T 是仿射变换当且仅当 $\gamma=(1-t)\boldsymbol{\alpha}+t\boldsymbol{\beta}$ 总可推出 $\gamma T=(1-t)\boldsymbol{\alpha}T+t(\boldsymbol{\beta}T)$.

16. 证明: 如果仿射子空间 M 是由 $m+1$ 个仿射无关的点 $\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_m$ 张成, 那么 M 同一个 m 维向量子空间平行.

17. 根据定义, F^n 中的超平面是 $n-1$ 维仿射子空间.

    (a) 证明: 坐标满足线性方程 $a_1x_1+\cdots+a_nx_n=c$ 的所有向量 $\boldsymbol{\xi}$ 组成的集合是一个超平面, 这里假定 $a,\cdots,a_n$ 不全为零.

    (b) 反过来证明: 每个超平面都满足这样的方程.

    (c) 求通过四个点 (1,0,1,0), (0,1,0,1), (0,1,1,0), (1,0,0,1) 的超平面方程.

18. 设 $\alpha_0,\cdots,\alpha_n$ 是 n 维向量空间 V 的 $n+1$ 个仿射无关的点, 且设 $\boldsymbol{\beta}_0,\cdots,\boldsymbol{\beta}_n$ 是 V 中任意 $n+1$ 个点. 证明: 存在一个且只存在一个 V 的仿射变换把每个 $\boldsymbol{\alpha}_i$ 变到 $\boldsymbol{\beta}_i$.

19. 证明: 如果仿射子空间 M 是由 $m+1$ 个仿射无关的点 $\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_m$ 张成, 并且由 $r+1$ 个仿射无关的点 $\boldsymbol{\beta}_0,\cdots,\boldsymbol{\beta}_r$ 张成, 那么 $m=r$.

*9.14 射 影 几 何

在实仿射平面中, 任意两个点在唯一的一条直线上, 任意两条不平行的直线交于唯一的一点. 我们现在来构造*实射影平面*, 在这个平面上,

(i) 任意两个不同的点在唯一的一条直线上.

(ii) 任意两条不同的直线交于唯一的一点.

相关联的性质 (i) 和 (ii) 显然在下述意义下彼此对偶, 即互换 "点" 和 "直线" 二词, 并在术语上稍加变化, 就可把性质 (i) 变为性质 (ii), 也可把性质 (ii) 变为性质 (i).

构造实射影平面 $P_2=P_2(\mathbf{R})$ 的一种方法如下所述. 取实数域 $\mathbf{R}$上一个三维向量空间 V_3, 并把 V_3 的一维向量子空间 (不是仿射子空间)S 称为 P_2 的一个 "点", 把 V_3 的二维子空间 L 称为 P_2 的 "直线". 进一步, 我们说点 S 在直线 L 上当且仅当子空间 S 包含在子空间 L 中.

我们来证明 $P_2(\mathbf{R})$ 的 "点" 和 "直线" 满足 (i) 和 (ii). 如果 S_1 和 S_2 分别是由向量 $\boldsymbol{\alpha}_1$ 和 $\boldsymbol{\alpha}_2$ 张成的一维子空间, 那么 $S_1\neq S_2$ 当且仅当 $\boldsymbol{\alpha}_1$ 和 $\boldsymbol{\alpha}_2$ 是线性无关. 于是 S_1 和 S_2 所在的唯一直线 L 就是由 $\boldsymbol{\alpha}_1$ 和 $\boldsymbol{\alpha}_2$ 张成的二维向量子空间, 这就证明了性质 (i). 其次, 如果直线 (二维子空间)L_1 和 L_2 是不同的, 那么子空间 L_1+L_2(是 L_1 与 L_2 的线性和) 必具有较高的维数, 于是它是整个三维空间 V_3. 因此根据 7.8 节定理 17, 有

$$\dim(L_1\cap L_2)=\dim L_1+\dim L_2-\dim(L_1+L_2)=2+2-3=1,$$

所以一维子空间 $L_1\cap L_2$ 是同时位于 L_1 和 L_2 上的唯一的一点. 这就证明了性质 (ii).

为了在射影平面 $P_2=P_2(\mathbf{R})$ 中得到适当的射影坐标, 取 V_3 是由所有 3- 实数组 (x_1,x_2,x_3) 组成的空间 $\mathbf{R}^3$. 那么每个非零的 3- 数组 (x_1,x_2,x_3) 确定 P_2 的一点 S; 当 $c\neq 0$ 时, 3- 数组 (x_1,x_2,x_3) 和 (cx_1,cx_2,cx_3) 确定同一个点 S. 我们把这些 3- 数组等同起来

$$(x_1,x_2,x_3)=(cx_1,cx_2,cx_3),\quad c\neq 0,$$

并称这些 3- 数组为点 S 的*齐次坐标*. 因为 V_3 的任意二维子空间 L 可以描述为一个齐次线性方程的解向量组成的集合, 所以 P_2 的直线 L 是那些齐次坐标满足方程

$$a_1x_1+a_2x_2+a_3x_3=0,\quad (a_1,a_2,a_3)\neq(0,0,0) \tag{58}$$

的点的轨迹. 我们可称 (a_1,a_2,a_3) 为直线 L 的*齐次坐标*. 显然, 当 $c\neq 0$ 时, 坐标 (a_1,a_2,a_3) 和 (ca_1,ca_2,ca_3) 确定同一条直线.

实射影平面有非常简单的几何表示. 点 S 的齐次坐标 (x_1,x_2,x_3), 通过乘上 $(x_1^2+x_2^2+x_3^2)^{-\frac{1}{2}}$ 可以标准化, 于是新坐标 (y_1,y_2,y_3) 满足 $y_1^2+y_2^2+y_3^2=1$, 并在单位球面上, 在这个球面上的两个对径点 (y_1,y_2,y_3) 和 $(-y_1,-y_2,-y_3)$ 确定 P_2 的同一点. 换句话说, P_2 的点可以通过把直径的两端点 (在球面上) 看作一个点而得到. 因为 V_3 的任意二维向量子空间 L 沿一个大圆截割单位球, 所以我们可以说, P_2 的一条直线是由单位球的一个大圆上的全体对径点对组成. 于是这又表明, 两条射影直线 (两个大圆) 交于一个射影点 (球面上一对对径点).

可以用同样的方法在任意域 F 上定义一个射影平面 $P_2(F)$. 在任何情形下, 显然有每个一维向量子空间 (cx_1,cx_2,cx_3)(其中 $x_3\neq 0$) 与仿射平面 $x_3=1$ 恰恰交于一点 $\left(\dfrac{x_1}{x_3},\dfrac{x_2}{x_3},1\right)$; 这个比 $\left(\dfrac{x_1}{x_3},\dfrac{x_2}{x_3}\right)$ 称为射影点 (cx_1,cx_2,cx_3) 的*非齐次坐标*. 但是 $x_3=0$ 的轨迹是一条射影直线, 它称为"*在无穷远处的直线*". 可以验证, 射影平面 P_2 的每条直线

$$L: a_1x_1+a_2x_2+a_3x_3=0$$

或者是"在无穷远处的直线"(当 $a_1=a_2=0$), 或者是仿射平面的一条直线 $a_1\left(\dfrac{x_1}{x_3}\right)+a_2\left(\dfrac{x_2}{x_3}\right)+a_3=0$, 加上"在无穷远处的直线"上的一个点 $(a_2,-a_1,0)$.

在任意域 F 上可以构造 n 维射影空间 P. 本质的一步是从高一维的向量空间 $V=F^{n+1}$ 开始. 那么 $P=P_n(F)$ 可以描述如下: P 的点是 V 的一维子空间 S, P 的 m *维子空间*是由位于 V 的某个 $m+1$ 维向量子空间 L 的所有点 S(P 的点) 组成的集合. 显然, 每个这样的子空间本身与由这个 $m+1$ 维向量空间 L 按同样方法确定的 m 维射影空间 P_m 同构. 如果 V 表示为 F 的 $(n+1)$- 元素组构成的空间 (用对于给定基的坐标表示), 那么 P_n 的每个点 S 可以用 $n+1$ 个齐次坐标 $(x_1,\cdots,x_{n+1})$ 给出, 并且坐标 $(cx_1,\cdots,cx_{n+1})$(其中 $c\neq 0$) 与坐标 $(x_1,\cdots,x_{n+1})$ 确定同一个点.

$P=P_n(F)$ 中的超平面 ($n-1$ 维子空间) 仍是由一个齐次方程

$$a_1x_1+\cdots+a_{n+1}x_{n+1}=0,\quad (a_1,\cdots,a_{n+1})\neq(0,\cdots,0) \tag{59}$$

给出的轨迹. 数组 $(a_1,\cdots,a_{n+1})$ 可以看作超平面的齐次坐标; 射影空间 P 和它的对偶射影空间 (这个空间的点是 P 的超平面) 之间的关系, 同向量空间 V 和它的对偶空间 V^* 之间的关系一样. 根据 7.7 节定理 13(关于齐次线性方程组解集合的维数的定理), 我们得到, r 个像 (59) 那样的线性无关的方程组可以确定一个 $n-r$ 维射影子空间.

设 $T:V\to V$ 是一个非奇异线性变换. 我们知道 (8.6 节定理 10 的推论 2), T 把 V 的每个一维子空间 S 变换到 V 的一个一维子空间 S^*. 因此 T 诱导出射影空

间 P 的点的变换 $S \longmapsto S^* = ST^*$, 这个变换 T^* 把射影子空间变换到射影子空间, 其维数保持不变. 我们称 T^* 是 P 的*射影变换*. 如果 T_1 和 T_2 是 V 的两个这样的线性变换, 那么乘积 T_1T_2 诱导出 P 上的一个变换 $(T_1T_2)^*$, 它是这两个诱导出的变换的乘积 $T_1^*T_2^*$. 因此所有射影变换组成的集合构成一个群, 即 n *维射影群*, 并且对应 $T \longmapsto T^*$ 是 $n+1$ 维全线性群到域 F 上 n 维射影群上的一个同态.

相对于 V 中给定的坐标系, 线性变换 T 是由一个非奇异 $(n+1)\times(n+1)$ 矩阵 (a_{ij}) 确定的. 那么变换 T^* 把具有齐次坐标 $(x_1,\cdots,x_{n+1})$ 的点变换到具有齐次坐标 $(y_1,\cdots,y_{n+1})$ 的点, 其中

$$y_j = x_1a_{1j} + \cdots + x_{n+1}a_{n+1,j} \quad (j = 1,\cdots,n+1). \tag{60}$$

定理 33 *$(n+1)\times(n+1)$ 矩阵 $\boldsymbol{A}$ 确定 P_n 的恒等射影变换 T^* 当且仅当 $\boldsymbol{A}$ 是单位矩阵 $\boldsymbol{I}$ 的数乘积 $c\boldsymbol{I}$(其中 $c \neq 0$).*

证明 在公式 (60) 中, 如果 $\boldsymbol{A} = c\boldsymbol{I}$, 则 $y_j = cx_j$: 即齐次坐标 $(x_1,\cdots,x_{n+1})$ 和 $(cx_1,\cdots,cx_{n+1})$ 确定 P 的同一个点, 所以 T^* 的确是恒等变换. 反过来, 假设 T^* 是恒等变换, 那么 T 一定把 $n+1$ 个单位向量 $\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_{n+1}$ 的每一个 $\boldsymbol{\varepsilon}_i$ 变换到某个数乘积 $c_i\boldsymbol{\varepsilon}_i$, 因此 $\boldsymbol{A}$ 一定是对角矩阵, 其对角线元素为 $c_1,\cdots,c_{n+1}$. 但是 T 也把向量 $(1,1,\cdots,1)$ 变换到它的一个数乘积, 而 $\boldsymbol{A}$ 把这个向量变换到 $(c_1,\cdots,c_{n+1})$. $(c_1,\cdots,c_{n+1})$ 是 $(1,1,\cdots,1)$ 的数乘积当且仅当 $c_1,\cdots,c_{n+1}$ 都相等. 因此 $\boldsymbol{A}$ 的确是 $\boldsymbol{I}$ 的一个数乘积.

推论 *域 F 上的 n 维射影群与 $n+1$ 维全线性群对于由恒等变换的非零数乘积组成的子群的商群同构.*

证明 映射 $T \longmapsto T^*$ 是全线性群到射影群中的同态; 定理 33 断言, 这个同态的核恰恰是由恒等变换的数乘积组成的集合. 因此根据 7.13 节定理 28 得此结论.

还可以推出, 两个矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{A}_1$ 确定同一个射影变换当且仅当 $\boldsymbol{A}_1 = c\boldsymbol{A}$($c$ 为某一标量).

对于一维射影直线, 射影变换具有形式

$$y_1 = ax_1 + bx_2, \quad y_2 = cx_1 + dx_2, \quad ad \neq bc. \tag{61}$$

按照非齐次坐标 $z = \dfrac{x_1}{x_2}$, $w = \dfrac{y_1}{y_2}$, 这个变换可写成线性分式变换

$$w = \frac{az+b}{cz+d}, \tag{62}$$

它是由 (61) 的第一个方程除以第二个方程而得到的. 公式 (62) 可以解释如下: 如果 $c = 0$, 则 (62) 把点 $z = \infty$ 变到点 $w = \infty$; 如果 $c \neq 0$, 则 (62) 把点 $z = \infty$ 变到点 $w = \dfrac{a}{c}$, 把点 $z = -\dfrac{d}{c}$ 变到点 $w = \infty$. 这些解释的正确性可以通过代回齐次坐标

并利用 (61) 式来验证. 在 n 维的情况, 也可能有用线性分式变换表示射影变换的类似的表达式

$$w_i = \frac{z_1 a_{1i} + \cdots + z_n a_{ni} + a_{n+1,i}}{z_1 b_1 + \cdots + z_n b_n + b_{n+1}} \quad (b_j = a_{j,n+1};\ i = 1, \cdots, n). \tag{62'}$$

我们已经看到, $P_n(F)$ 的射影变换把直线变到直线. 反过来, 实射影空间 $P_n(\mathbf{R})$ 的任意把直线变到直线的一一变换, 当 $n \geqslant 2$ 时, 是射影变换 (见习题 6), 这是一个经典的结果.

三个变量的二次型确定射影平面中的一个轨迹

$$\sum_{i,j} x_i b_{ij} x_j = 0 \quad (i, j = 1, 2, 3), \tag{63}$$

这是因为如果坐标 (x_1, x_2, x_3) 满足这个方程, 那么任意 "数乘" 积 (cx_1, cx_2, cx_3) 也满足这个方程. 这个轨迹称为*射影二次曲线*; 这个二次曲线的 (射影)秩是系数矩阵 $\boldsymbol{B}$ 的秩. 如果排除 "在无穷远外的直线", 这个射影二次曲线 (63) 就变成普通二次曲线. 在实射影平面中, 任意非退化二次曲线 (即椭圆、双曲线或抛物线), 根据 9.9 节, 它与下列四个方程中的一个所确定的曲线等价:

$$x_1^2 + x_2^2 + x_3^2 = 0, \quad x_1^2 + x_2^2 - x_3^2 = 0, \tag{64}$$

$$-x_1^2 - x_2^2 - x_3^2 = 0, \quad x_1^2 - x_2^2 - x_3^2 = 0. \tag{64'}$$

这些方程中, 等号左边各项符号改变并不改变其轨迹, 因此由 (64′) 给出的二次曲线本质上就是 (64) 给出的二次曲线. (64) 第一条二次曲线是空的. 因此我们得出, *任意两个非退化二次曲线在实射影平面中是射影等价的*.

习　　题

1. 在域 F 上的三维射影空间中, 证明:
   (a) 任意两个不同的点在一条且只在一条直线上.
   (b) 任意不在一直线上的三个点在一个且只在一个平面上.
2. 推广习题 1 到 n 维射影空间.
3. 在域 $\mathbf{Z}_2$ 上的射影平面中列出所有的点和直线, 以及每条直线上的全部点.
4. 在含有 n 个元素的有限域上的射影平面中, 证明: 存在 n^2+n+1 个点, n^2+n+1 条直线, 并且每条直线上有 $n+1$ 个点.
5. 四个不同数 z_1, z_2, z_3, z_4 的交比定义为比值 $\dfrac{(z_3-z_1)(z_4-z_2)}{(z_3-z_2)(z_4-z_1)}$(当 z_i 中有一个是 ∞ 时, 用适当的约定). 证明: 交比是任意线性分式变换 (62) 之下的不变量.
6. 证明: 在复射影平面上, 变换 $(z_1, z_2, z_3) \longmapsto (z_1^*, z_2^*, z_3^*)$ 把直线变到直线, 但这个变换不是射影变换. (星号表示复共轭.)

7. 如果去掉 "在无穷远外的直线" $x_3=0$, 那么射影二次曲线 $x_1^2=2x_2x_3$ 在仿射平面中表示什么?
8. (a) 证明: 每个非退化实二次曲面与一个球面或与一个单叶双曲面射影等价.

   (b) 椭圆抛物面与上面哪个曲面射影等价? 双曲抛物面与上面哪个曲面射影等价?

   (c) 证明: 球面与单叶双曲面不射影等价.
9. 证明: 在射影直线中给出任意两个不同点的 3- 数组 (z_1,z_2,z_3) 和 (w_1,w_2,w_3), 则存在一个射影变换 (62), 它把每个 z_i 变到相应的 w_i.
10. 设 (p_1,p_2,p_3,p_4) 和 (q_1,q_2,q_3,q_4) 是射影平面中的任意两个点的 4- 数组. 证明: 存在一个射影变换 (62′), 它把每个 p_i 变到相应的 q_i.

第 10 章 行列式与标准型

10.1 行列式的定义和基本性质

在任意域上每个方阵 $\boldsymbol{A}$ 都有一个行列式; 虽然行列式能够用来研究矩阵的秩和求解联立线性方程组, 但是它在矩阵论中最重要的应用是定义矩阵的特征多项式. 在这一章中, 我们来定义行列式, 研究它的几何性质, 并指出矩阵 $\boldsymbol{A}$ 的特征多项式和矩阵的特征根 (特征值) 之间的关系. 然后, 用这些概念来研究矩阵在相似变换之下的标准型.

联立线性方程组的求解公式自然导出行列式. 两个线性方程

$$a_1x + b_1y = k_1,$$
$$a_2x + b_2y = k_2$$

在 $a_1b_2 - a_2b_1 \neq 0$ 的假定下, 有唯一解

$$x = \frac{k_1b_2 - k_2b_1}{a_1b_2 - a_2b_1}, \quad y = \frac{a_1k_2 - a_2k_1}{a_1b_2 - a_2b_1}.$$

出现在分子和分母中的多项式称为*行列式*,

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1, \quad \begin{vmatrix} k_1 & b_1 \\ k_2 & b_2 \end{vmatrix} = k_1b_2 - k_2b_1. \tag{1}$$

类似地, 我们可以计算三个联立线性方程组 $\Sigma a_{ij}x_j = k_i$ 的解. 结果每个解 x_j 的分母是

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{aligned} &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad -a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}, \end{aligned} \tag{2}$$

等式的右边是六项乘积的代数和. 每一项包含矩阵第一行的某元素 a_{1i}, 同时也包含第二行的某元素和第三行的某元素. 每一项还可以表示成包含着不同列的元素乘积, 所以 (2) 中的每一项可写成 $a_{1-}a_{2-}a_{3-}$, 空白的地方添上列标 1,2,3 的某一置换. 在六个可能的置换中, 三个偶置换 $I, (123), (132)$ 出现在带正号的乘积项中, 而三个奇置换出现在带负号的乘积项中. 经验表明, n 个未知数的 n 个线性方程的解可以表示为类似的公式.

定义　$n\times n$ 矩阵 $\boldsymbol{A}=(a_{ij})$ 的行列式是元素[①] $a_{ij}=a(i,j)$ 的多项式:

$$\det \boldsymbol{A}=|\boldsymbol{A}|=\sum_{\phi}\operatorname{sgn}\phi\left[\prod_{i=1}^{n}a_{i,i\phi}\right]$$
$$=\sum_{\phi}(\text{sng }\phi)a(1,1\phi)a(2,2\phi)\cdots a(n,n\phi). \tag{3}$$

这里是对整数 $1,\cdots,n$ 的所有 $n!$ 个不同的置换 ϕ 求和. 每个乘积项 $\Pi a_{i,i\phi}$ 前面的因子 $\operatorname{sgn}\phi$ 是 $+1$ 还是 -1, 根据 ϕ 是偶置换还是奇置换而定.

于是行列式 $|\boldsymbol{A}|=\det(a_{ij})$ 是 $n!$ 项 $\pm a_{1-}a_{2-}\cdots a_{n-}$ 的和, 这里空白处添上数字 $1,\cdots,n$ 的各种置换 ϕ. 把 a_{ij} 写成 $a(i,j)$, 设 $i\phi$ 是 i 在 ϕ 之下的像, 于是一般项可写成 $\pm a(1,1\phi)a(2,2\phi)\cdots a(n,n\phi)$, 这里正负号 "±" 称为 $\operatorname{sgn}\phi$(signum ϕ 的缩写). 对于和中的每个乘积项, 恰好含有矩阵每一行中一个元素, 并且恰好含有矩阵每一列中一个元素.

矩阵的每一行在 $|\boldsymbol{A}|$ 的每一项中出现一次且只出现一次, 这就意味着 $|\boldsymbol{A}|$ 是 $\boldsymbol{A}$ 的第 i 行元素 $a_{i1},\cdots,a_{in}$ 的线性齐次函数. 把每个这样的 a_{ij} 的系数合并起来, 我们就得到表达式

$$|\boldsymbol{A}|=\boldsymbol{A}_{i1}a_{i1}+\boldsymbol{A}_{i2}a_{i2}+\cdots+\boldsymbol{A}_{in}a_{in}, \tag{4}$$

这里 a_{ij} 的系数 $\boldsymbol{A}_{ij}$ 称为 a_{ij} 的余子式(即代数余子式); 它是 $\boldsymbol{A}$ 中划去第 i 行剩下的各行元素的多项式. 这个余子式还可以描述成偏导数 $\boldsymbol{A}_{ij}=\dfrac{\partial|\boldsymbol{A}|}{\partial a_{ij}}$. 因为 $|\boldsymbol{A}|$ 的每一项只包含每一行和每一列中一个元素, 所以余子式 $\boldsymbol{A}_{ij}$ 即不包含第 i 行的元素也不包含第 j 列的元素. 它只包含 "子式" 或子矩阵$\boldsymbol{M}_{ij}$ 的元素, 子矩阵 $\boldsymbol{M}_{ij}$ 是从矩阵 $\boldsymbol{A}$ 中划去第 i 行和第 j 列的所有元素所得到的矩阵.

$|\boldsymbol{A}|$ 对于行和列是对称的.

定理 1　设 $\boldsymbol{A}^{\mathrm{T}}$ 表示 $\boldsymbol{A}$ 的转置矩阵, 则 $|\boldsymbol{A}^{\mathrm{T}}|=|\boldsymbol{A}|$.

证明　$\boldsymbol{A}^{\mathrm{T}}$ 的元素 $a_{ij}^{\mathrm{T}}=a_{ji}$ 是把 $\boldsymbol{A}$ 的元素 a_{ij} 的下标颠倒过来而得到的. 当 $j=i\phi$, $i=j\phi^{-1}$, $|\boldsymbol{A}|$ 的一般项是

$$(\operatorname{sgn}\phi)\prod_{i}a(i,i\phi)=(\operatorname{sgn}\phi)\prod_{j}a(j\phi^{-1},j)=(\operatorname{sgn}\phi)\prod_{j}a^{\mathrm{T}}(j,j\phi^{-1}).$$

它也是 $|\boldsymbol{A}^{\mathrm{T}}|$ 的一般项, 因为每一置换是某个置换 ϕ 的逆置换 ϕ^{-1}, 而且正负号也相同, $\operatorname{sgn}\phi=\operatorname{sgn}\phi^{-1}$, 这是根据 ϕ 是偶置换 (即在交错群中) 当且仅当 ϕ 的逆 ϕ^{-1} 也是偶置换 (6.10 节). 因此 $|\boldsymbol{A}|=|\boldsymbol{A}^{\mathrm{T}}|$. 　证毕

对行列式进行初等行运算时, 将会产生什么样的效果呢?

① 这里指的元素是域 F 中的元素, 或更一般地是指交换环中的元素.

法则 1 矩阵 $\boldsymbol{A}$ 的第 i 行乘上标量 $c \neq 0$, 则相应的行列式就乘上 c. 这是因为在线性齐次表达式 (4) 中, 对第 i 行的每个元素 $a_{i1}, \cdots, a_{in}$ 乘上一个因子 c 就意味着对 $|\boldsymbol{A}|$ 乘上同一个因子 c.

法则 2 矩阵 $\boldsymbol{A}$ 的两行交换, 行列式 $|\boldsymbol{A}|$ 变号. 根据对称性 (定理 1), 我们可以改为证明矩阵的两列交换, 行列式变号. 这种交换可以用列标的奇置换 ϕ_0 来表示, 于是可用矩阵 $\boldsymbol{B} = (b_{ij})$ 来代替 $\boldsymbol{A}$, 其中 $b_{ij} = b(i,j) = a(i, j\phi_0)$, 那么

$$|\boldsymbol{B}| = \sum_{\phi} (\text{sgn } \phi) \prod_i b(i, i\phi) = \sum_{\phi} (\text{sgn } \phi) \prod_i a(i, i\phi\phi_0).$$

因为全体置换构成一个群, 所以全体乘积 $\phi\phi_0$(这里 ϕ_0 是固定的, ϕ 跑遍所有置换) 包含全部置换, 因此上面的 $|\boldsymbol{B}|$ 含有 $|\boldsymbol{A}|$ 的所有项. 只是每一项的正负号改变了, 这是因为 ϕ_0 是奇置换, 所以当 ϕ 是奇置换时, $\phi\phi_0$ 是偶置换, 反过来, 当 ϕ 是偶置换时, $\phi\phi_0$ 是奇置换, 于是 $\text{sgn } \phi\phi_0 = -\text{sgn } \phi$. 这就证明了法则 2.

引理 1 如果矩阵 $\boldsymbol{A}$ 有相同的两行, 那么 $|\boldsymbol{A}| = 0$.

证明 根据定理 1, 只须证明当 $\boldsymbol{A}$ 有相同的两列时, $|\boldsymbol{A}| = 0$. 设 ψ 是把两个相同的列进行交换的对换, 那么表达式 (3) 中的所有被加项 $(\text{sgn}\phi)\Pi a(i, i\phi)$ 是按 $\{\phi, \psi\phi\}$ 成对出现, 而 $\{\phi, \psi\phi\}$ 是由 ψ 生成的二元素子群的陪集组成. 因为 ψ 是奇置换, 所以 $\text{sgn } \phi = -\text{sgn } \psi\phi$; 又因为两个列是相同的, 所以有 $\Pi a(i, i\phi) = \Pi a(i, i\psi\phi)$. 因此成对的被加项其数值相等符号相反, 所以它们的和是零. 证毕

为了考虑伴随矩阵 (10.2 节), 用一个方程表示引理 1 是方便的. 在 $\boldsymbol{A}$ 中, 用第 k 行来代替第 i 行, 则两行变成一样了, 其行列式等于零. 但是这个行列式还可以在线性齐次表达式 (4) 中用第 k 行代替第 i 行而得到, 于是

$$0 = A_{i1}a_{k1} + A_{i2}a_{k2} + \cdots + A_{in}a_{kn} \quad (i \neq k). \tag{5}$$

法则 3 矩阵第 k 行乘上常数 c 加到第 i 行上, 其行列式保持不变. 这个运算是用 $a_{ij} + ca_{kj}$ 代替每个 a_{ij}; 根据线性齐次表达式 (4), 新的行列式是

$$\sum_j \boldsymbol{A}_{ij}(a_{ij} + ca_{kj}) = \sum_j \boldsymbol{A}_{ij}a_{ij} + c\sum_j \boldsymbol{A}_{ij}a_{kj} = |A| + 0,$$

上式最后一个等式是根据 (4) 和 (5) 得到的. 行列式的确没变.

这些法则可以用初等矩阵加以概括. 任意初等行运算把单位矩阵 $\boldsymbol{I}$ 变成初等矩阵 $\boldsymbol{E}$, 而把 $\boldsymbol{A}$ 变成乘积 $\boldsymbol{EA}$. 行列式 $|\boldsymbol{I}| = 1$ 因而变成 $|\boldsymbol{E}| = c, -1$ 或 1(对应于法则 1,2 或 3), 而 $|\boldsymbol{A}|$ 变成 $|\boldsymbol{EA}| = c|\boldsymbol{A}|, -|\boldsymbol{A}|$ 或 $|\boldsymbol{A}|$(对应于相应的法则). 这就证明了 $|\boldsymbol{EA}| = |\boldsymbol{E}||\boldsymbol{A}|$. 根据对称性 (定理 1), 同样可应用于右乘因子 $\boldsymbol{E}$. 这就建立了

定理 2 如果 $\boldsymbol{E}$ 是初等矩阵, 那么

$$|\boldsymbol{EA}| = |\boldsymbol{E}||\boldsymbol{A}| = |\boldsymbol{AE}|.$$

另一个法则是, 从上面讨论的子矩阵 $\boldsymbol{M}_{ij}$ 可以明显地得到余子式的表达式.

法则 4 $\boldsymbol{A}_{ij} = (-1)^{i+j}|\boldsymbol{M}_{ij}|$, 口头上说就是, 每个余子式 $\boldsymbol{A}_{ij}$ 是由相应的子矩阵的行列式前面添上符号 $(-1)^{i+j}$ 而得到. 正负号 $(-1)^{i+j}$ 可以在西洋跳棋盘 (想象在方格盘中 $+, -$ 号相间排列, 并且左上角第一格中是 $+$ 号) 上的 (i,j) 位置上得到. 首先我们对 $i=j=1$ 来证明这个法则. 从定义 (3) 立即看出, 包含 a_{11} 的项恰恰是满足条件 $1\phi = 1$ 的那些置换 ϕ 所对应的项. 这种类型的偶置换 (或奇置换) 实际上是剩下的数字 $2, \cdots, n$ 的偶置换 (或奇置换), 所以这些项划去 a_{11} 后恰恰就是 $|\boldsymbol{M}_{11}|$ 的展开式的所有项. 任意其他余子式 $\boldsymbol{A}_{ij}$ 可通过把 a_{ij} 移到左上角而化成上述特殊情形, 而 a_{ij} 移到左上角位置是相继进行 $i-1$ 次相邻两行的交换和 $j-1$ 次相邻两列的交换得到的. 这些初等运算并不改变 $|\boldsymbol{M}_{ij}|$, 因为它不影响 $\boldsymbol{M}_{ij}$ 中行与列的相互位置, 但是这些运算却改变了 $|\boldsymbol{A}|$ 的符号, 因此 a_{ij} 的余子式的符号改变了 $i+j-1-1$ 次. 通过这样的简化就证明了法则 4.

一个特别有用的情形是, 矩阵的第一行除了第一个元素外其他所有元素都是零. 那么表达式 (4) 只须包含第一个余子式 $|\boldsymbol{M}_{11}| = \boldsymbol{A}_{11}$, 所以

$$\begin{vmatrix} c & \boldsymbol{O} \\ \boldsymbol{K} & \boldsymbol{B} \end{vmatrix} = c|\boldsymbol{B}|, \tag{6}$$

这里 $\boldsymbol{O}$ 是 $1\times(n-1)$ 矩阵, $\boldsymbol{K}$ 是 $(n-1)\times 1$ 矩阵, $\boldsymbol{B}$ 是 $(n-1)\times(n-1)$ 矩阵. 根据这个法则并用归纳法, 我们得到下面结果.

引理 2 *三角形矩阵的行列式等于它的对角线元素之积.*

上述各法则提供了一系列计算行列式 $|\boldsymbol{A}|$ 的方法. 通过初等运算把矩阵 $\boldsymbol{A}$ 化为三角形矩阵 $\boldsymbol{T}$, 用 t 表示所用行 (或列) 交换的次数, 用 $c_1, \cdots, c_s$ 表示乘到 $\boldsymbol{A}$ 的行 (或列) 上的各个标量. 根据定理 2, $|\boldsymbol{A}| = (-1)^t(c_1\cdots c_s)^{-1}|\boldsymbol{T}|$, 再用引理 2 设 $|\boldsymbol{T}| = t_{11}\cdots t_{nn}$, 这样就计算出行列式的值.

习 题

1. 从行列式的定义直接证明引理 2.
2. 计算 7.6 节习题 2 中各矩阵的行列式.
3. (a) 设 $\boldsymbol{A} = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 2 & 1 & 1 \end{pmatrix}$, 分别用第一行的子式或第一列的子式计算行列式 $|\boldsymbol{A}|$, 并比较这两个计算结果.

   (b) 假定上面矩阵 $\boldsymbol{A}$ 的元素是模 2 整数, 计算 $|\boldsymbol{A}|$.
4. 写出一般 4×4 行列式展开式中的所有正项.
5. 设 n 是奇数, 并且 $1+1\neq 0$, 证明: $n\times n$ 斜对称矩阵 $\boldsymbol{A}$ 的行列式为零.

6. (a) 推导下面范得蒙(Vandermonde) 行列式的展开式

$$\begin{vmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{vmatrix} = (x_2-x_1)(x_3-x_1)(x_3-x_2).$$

(b) 推广这个结果到 4×4 的情形.

(c) 推广这个结果到 $n\times n$ 的情形, 即证明: 如果 $a_{ij}=x_i^{j-1}$, 那么 $|\boldsymbol{A}|=\prod\limits_{i>j}(x_i-x_j)$.

7. 证明: 对任意 4×4 斜对称矩阵 $\boldsymbol{A}$, 有 $|\boldsymbol{A}|=(a_{12}a_{34}-a_{13}a_{24}+a_{14}a_{23})^2$.

8. (a) 证明: 任意置换矩阵的行列式等于 ±1.

(b) 证明: 单项矩阵的行列式等于全体非零元素的乘积再乘上 ±1.

9. 如果 $n\times n$ 实矩阵 $\boldsymbol{A}$, 对于 $i=1,\cdots,n$, 有 $\sum\limits_{j\neq i}|a_{ij}|<a_{ii}$, 则称 $\boldsymbol{A}$ 为对角优势矩阵. 证明: 如果 $\boldsymbol{A}$ 是对角优势矩阵, 那么 $|\boldsymbol{A}|>0$.

10. 证明: 平面上两个点 (a_1,a_2) 和 (b_1,b_2) 联线的方程是

$$\begin{vmatrix} x_1 & x_2 & 1 \\ a_1 & a_2 & 1 \\ b_1 & b_2 & 1 \end{vmatrix} = 0.$$

*11. (a) 证明: 如果矩阵 $\boldsymbol{A}$ 的每个元素 a_{ij} 是 x 的函数, 那么 $\dfrac{\mathrm{d}|\boldsymbol{A}|}{\mathrm{d}x}=\sum\limits_{j,k=1}^{n}\dfrac{\mathrm{d}a_{jk}}{\mathrm{d}x}\boldsymbol{A}_{jk}$.

(b) 用此验证 $\boldsymbol{A}_{ij}=\dfrac{\partial|\boldsymbol{A}|}{\partial a_{ij}}$.

*12. 证明: 如果 $\boldsymbol{A}$ 和 $\boldsymbol{C}$ 是方阵, 那么有

$$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{C} \end{vmatrix} = |\boldsymbol{A}||\boldsymbol{C}|.$$

*13. 设 $\boldsymbol{\Omega}$ 是 $n\times n$ 矩阵 (ω^{ij}), 这里 ω 是 n 次本原复单位根, 证明: 如果 $n\equiv1(\bmod 4)$, 那么有 $|\boldsymbol{\Omega}|=n^{\frac{n}{2}}$.

10.2 行列式的乘积

在初等行运算与初等列运算下, 任意方阵 $\boldsymbol{A}$ 等价于对角矩阵 $\boldsymbol{D}$ (8.9 节定理 18), 所以像 8.8 节定理 13 那样, 矩阵 $\boldsymbol{A}$ 可以由矩阵 $\boldsymbol{D}$ 左乘和右乘一些初等矩阵 $\boldsymbol{E}_i$ 和 $\boldsymbol{E}^{(i)}$ 得到, 即

$$\boldsymbol{A}=\boldsymbol{E}_s\cdots\boldsymbol{E}_1\boldsymbol{D}\boldsymbol{E}^{(1)}\cdots\boldsymbol{E}^{(t)}. \tag{7}$$

根据定理 2 的法则 $|\boldsymbol{EA}|=|\boldsymbol{E}||\boldsymbol{A}|$ 和 $|\boldsymbol{AE}|=|\boldsymbol{A}||\boldsymbol{E}|$, 在乘积 (7) 的行列式中, 可以同时对每个因子 $\boldsymbol{E}$ 取行列式 $|\boldsymbol{E}|$, 于是得到

$$|\boldsymbol{A}|=|\boldsymbol{E}_s|\cdots|\boldsymbol{E}_1||\boldsymbol{D}||\boldsymbol{E}^{(1)}|\cdots|\boldsymbol{E}^{(t)}|. \tag{8}$$

因为每个 $|\boldsymbol{E}_i| \neq 0$, 所以整个行列式 $|\boldsymbol{A}| \neq 0$ 当且仅当 $|\boldsymbol{D}| \neq 0$. 这个标准型 $\boldsymbol{D}$ 在主对角线上恰恰有 r 个元素 1, 这里 r 是矩阵 $\boldsymbol{A}$ 的秩, 而行列式 $|\boldsymbol{D}|$ 是它的 n 个对角线元素之积. 因此 $|\boldsymbol{D}| \neq 0$ 当且仅当 $r = n$, 即当且仅当 $\boldsymbol{A}$ 是非奇异的. 所以由 (8) 式证明了

定理 3 方阵 $\boldsymbol{A}$ 是非奇异的当且仅当 $|\boldsymbol{A}| \neq 0$.

行列式的计算 公式 (8) 也为 $n \times n$ 行列式的数值计算提供了一个有效方法. 我们像高斯消去法那样进行计算, 逐个把对角线元素用 1 代替, 这样容易算出对角线元素的乘积; 因为除了某一行 (或列) 乘以常数的初等矩阵外, 其他所用的初等矩阵的行列式都是 ± 1, 所以这足以计算出整个行列式. 例如,

$$\begin{vmatrix} 2 & 3 & 4 & 1 \\ 4 & -1 & 2 & 3 \\ -6 & 5 & 2 & 6 \\ 8 & 5 & 7 & -2 \end{vmatrix} = 2\begin{vmatrix} 1 & \frac{3}{2} & 2 & \frac{1}{2} \\ 0 & -7 & -6 & 1 \\ 0 & 14 & 14 & 9 \\ 0 & -7 & -9 & -6 \end{vmatrix} = -14\begin{vmatrix} 1 & \frac{3}{2} & 2 & \frac{1}{2} \\ 0 & 1 & \frac{6}{7} & -\frac{1}{7} \\ 0 & 0 & 2 & 11 \\ 0 & 0 & -3 & -7 \end{vmatrix},$$

因此这个行列式等于 $(-14)(19) = -266$.

一个非奇异矩阵 $\boldsymbol{A}$ 是初等矩阵的乘积 $\boldsymbol{A} = \boldsymbol{E}_t \cdots \boldsymbol{E}_1$. 如果 $\boldsymbol{B} = \boldsymbol{E}_s^* \cdots \boldsymbol{E}_1^*$ 是另一个这样的矩阵, 那么乘积 $\boldsymbol{AB}$ 的行列式可像 (8) 式那样进行计算, 结果为

$$|\boldsymbol{AB}| = |\boldsymbol{E}_t \cdots \boldsymbol{E}_1 \boldsymbol{E}_s^* \cdots \boldsymbol{E}_1^*| = |\boldsymbol{E}_t| \cdots |\boldsymbol{E}_1| \cdot |\boldsymbol{E}_s^*| \cdots |\boldsymbol{E}_1^*| = |\boldsymbol{A}| \cdot |\boldsymbol{B}|$$

定理 4 乘积矩阵的行列式是各矩阵的行列式的乘积: $|\boldsymbol{AB}| = |\boldsymbol{A}| \cdot |\boldsymbol{B}|$.

证明 上述计算仅给出当 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是非奇异矩阵时, 这个法则的证明. 而当 $\boldsymbol{A}$ 或 $\boldsymbol{B}$ 是奇异矩阵时, 因此 $\boldsymbol{AB}$ 也是奇异矩阵, 所以 $|\boldsymbol{AB}| = |\boldsymbol{A}| \cdot |\boldsymbol{B}|$ 的两边都等于零.

证毕

行列式 $|\boldsymbol{A}| \neq 0$ 的矩阵 $\boldsymbol{A}$ 的逆存在, 并且可以用 $\boldsymbol{A}$ 的余子式明显地求出. 最初的包含余子式的方程 (4) 和 (5) 可以写成

$$a_{k1}\boldsymbol{A}_{i1} + \cdots + a_{kn}\boldsymbol{A}_{in} = \delta_{ki}|\boldsymbol{A}|, \quad \delta_{ki} = \begin{cases} 1, & \text{当} i = k, \\ 0, & \text{当} i \neq k. \end{cases} \tag{9}$$

δ_{ki} 这个数恰好是单位矩阵 $\boldsymbol{I} = (\delta_{ki})$ 的 (k, i) 位置上的元素. 方程 (9) 很像矩阵的乘积, 如果把余子式的下标交换, 那么 (9) 式左边就给出了矩阵 $\boldsymbol{A}$ 和余子式矩阵的转置矩阵的乘积的 (k, i) 位置上的元素. (9) 式右边是用标量 $|\boldsymbol{A}|$ 乘上单位矩阵后的 (k, i) 位置上的元素, 所以有

$$\boldsymbol{A}(\boldsymbol{A}_{ij})^{\mathrm{T}} = |\boldsymbol{A}|\boldsymbol{I}. \tag{10}$$

出现在这个方程中的矩阵 $(\boldsymbol{A}_{ij})^{\mathrm{T}}$ 是 $\boldsymbol{A}$ 的全体元素的余子式构成的矩阵的转置矩阵, 称为 $\boldsymbol{A}$ 的*伴随矩阵*. 在 $|\boldsymbol{A}| = 1$ 的情况下, 方程 (10) 表明, $\boldsymbol{A}$ 的伴随矩阵就是 $\boldsymbol{A}$ 的逆矩阵. 一般地, 如果 $|\boldsymbol{A}| \neq 0$, (10) 式证明了

定理 5 如果 $|A| \neq 0$, 那么矩阵 A 的逆是 $A^{-1} = |A|^{-1}(A_{ij})^{\mathrm{T}}$.

解 n 个未知数 n 个线性方程的克莱姆法则是这个逆矩阵公式的推论. 已知的线性方程组形为

$$\sum_j a_{ij}x_j = b_i,$$

其中 i 和 j 是从 1 到 n. 用矩阵表示, 这个方程组就是 $AX = B$(X 和 $B = (b_1, \cdots, b_n)^{\mathrm{T}}$ 都是 n 维列向量). 如果 A 是非奇异的, 那么用 A^{-1} 左乘这个方程得到唯一的向量解 $X = (x_1, \cdots, x_n)^{\mathrm{T}} = A^{-1}B$. 如果我们注意到逆矩阵 A^{-1} 的 (i,j) 位置上的元素刚好是 $\dfrac{A_{ji}}{|A|}$, 那么这个解可以展开表出. 这就证明了

定理 6(克莱姆法则) 如果 n 个未知数 n 个线性方程

$$\sum_j a_{ij}x_j = b_i$$

的系数矩阵 $A = (A_{ij})$ 是非奇异的, 那么方程组有唯一解

$$x_j = \frac{A_{1j}b_1 + \cdots + A_{nj}b_n}{|A|}, \quad j = 1, \cdots, n, \tag{11}$$

其中 A_{ij} 是系数矩阵 A 中元素 a_{ij} 的余子式.

这个公式的分子本身可以写成行列式, 因为它是用常数列 $(b_1, \cdots, b_n)^{\mathrm{T}}$ 代替矩阵 A 的第 j 列而得到的行列式按第 j 列的余子式的展开式. 可是对于大的联立线性方程组, 通过化矩阵 (或增广矩阵) 为行等价梯形矩阵 (7.7 节) 的方法求解, 通常更为有效.

显然, 克莱姆法则可以应用到任意域上, 特别可以应用到 2.3 节中所讨论的所有方程组 (参看后面的习题 9). 克莱姆法则对于求解二个或三个未知数的联立线性方程组是特别方便的.

附录 行列式与秩 长方矩阵 A 的子矩阵 (或 "子式") 是从 A 划去 A 的某些行和某些列而得到的任意矩阵 (这里包括一行也没划去或一列也没划去的情况). 任意长方矩阵 $A \neq O$ 的 "行列式秩"d 可以定义为行列式不为零的 A 的最大子式的行数, 换句话说, d 具有性质: (i)A 至少有一个 $d \times d$ 子式 M, 它的行列式 $|M| \neq 0$; (ii) 如果 $h > d$, 则 A 的每个 $h \times h$ 子式 N, 都有 $|N| = 0$. 可以证明, 任意矩阵的秩等于它的行列式秩.

习 题

1. 写出 2×2 矩阵 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 的伴随矩阵, 并计算 A 和它的伴随矩阵的乘积.

2. (a) 计算 7.6 节习题 2(a) 中的矩阵的伴随矩阵. 对于这种情形验证关于矩阵和它的伴随矩阵的乘积法则
   (b) 对 7.6 节习题 2(b) 做同样的问题.
3. 用伴随矩阵的方法, 求 8.8 节 (50) 式中的 4×4 初等矩阵 $\boldsymbol{H}_{24}, \boldsymbol{I}+2\boldsymbol{E}_{33}$ 和 $\boldsymbol{I}+d\boldsymbol{E}_{21}$ 的逆矩阵.
4. 用伴随矩阵的方法, 求 8.8 节习题 5 的矩阵的逆.
5. 证明：如果 $\boldsymbol{A}$ 是非奇异矩阵, 那么 $|\boldsymbol{A}^{-1}|=|\boldsymbol{A}|^{-1}$.
6. 证明：奇异矩阵和它的伴随矩阵的乘积是零矩阵.
7. 证明：任意正交矩阵的伴随矩阵是它的转置矩阵.
8. 写出三个未知数的三个线性方程的克莱姆法则.
9. 用克莱姆法则, 求解 2.3 节习题 1 的联立同余式.
10. (a) 证明：下面一对齐次线性方程

$$a_1x+b_1y+c_1z=0, \qquad a_2x+b_2y+c_2z=0$$

有一组解

$$x=\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}, \quad y=\begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix}, \quad z=\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

   (b) 什么时候这组解是整个解集合的基?
   (c) 对四个未知数的三个线性方程推导类似的公式.
11. 证明：正交矩阵的行列式是 ± 1.
12. 证明：矩阵 $\boldsymbol{A}$ 的伴随矩阵的行列是 $|\boldsymbol{A}|^{n-1}$.
13. 证明：矩阵 $\boldsymbol{A}$ 的伴随矩阵的伴随矩阵是 $|\boldsymbol{A}|^{n-2}\boldsymbol{A}$.
14. 直接由行列式秩的定义证明：初等行运算不改变行列式秩.

*15. (a) 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是 3×3 矩阵, 证明：$\boldsymbol{AB}$ 的任意 2×2 子矩阵的行列式是一些项的和, 其中每一项是 $\boldsymbol{A}$ 的一个 2×2 子矩阵的行列式与 $\boldsymbol{B}$ 的一个 2×2 子矩阵的行列式的乘积.
   (b) 推广这个结果, 并用它证明 $\operatorname{rank}(\boldsymbol{AB})\leqslant \operatorname{rank}\boldsymbol{A}$.

*16. 设 $n\times n$ 矩阵 $\boldsymbol{A}$ 的秩为 r, 证明：$\boldsymbol{A}$ 的伴随矩阵的秩 s 如下确定：若 $r=n$, 则 $s=n$; 若 $r=n-1$, 则 $s=1$; 若 $r<n-1$, 则 $s=0$.

*17. 证明：任意矩阵的秩等于它的行列式秩.

10.3 作为体积的行列式

$n\times n$ 实矩阵的行列式在几何上可以解释为 n 维欧几里得空间中的体积. 这是从平行四边形面积公式得到启发的.

每个以向量 $\boldsymbol{\alpha}_1$ 和 $\boldsymbol{\alpha}_2$ 为行的 2×2 实矩阵 $\boldsymbol{A}$ 表示一个以 $\boldsymbol{O}, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$ 为顶点的平行四边形. 反过来, 每个这样的平行四边形确定一个矩阵 (参看图 10-1).

这个平行四边形的面积是

$$底 \times 高 = |\boldsymbol{\alpha}_1||\boldsymbol{\alpha}_2||\sin C|, \tag{12}$$

式中 C 表示已知向量 $\boldsymbol{\alpha}_1$ 和 $\boldsymbol{\alpha}_2$ 之间的夹角. 根据 7.9 节的余弦公式 (41), 这个面积的平方等于

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1)(\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2)(1-\cos^2 C) = (\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1)(\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2) - (\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)(\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1).$$

这个结果看起来很像 2×2 矩阵的行列式, 实际上, 它就是矩阵 $((\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)) = \boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ 的行列式.

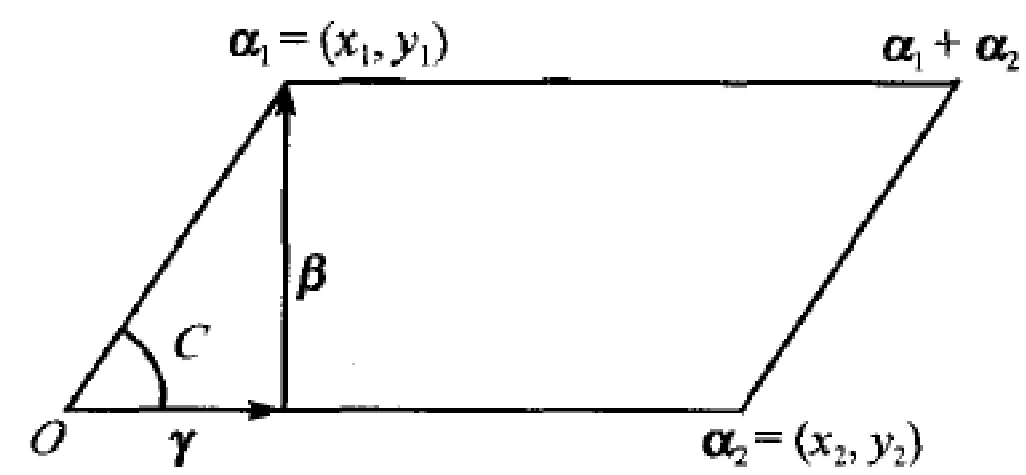


图 10-1

任意维欧几里得空间中的平行四边形都具有类似的公式, 甚至还可以推广到 n 维欧几里得空间中的 m 维类似于平行四边形的几何体, 这些几何体称为平行六面体.

为了进行这种推广, 设 $\boldsymbol{A}$ 是任意一个以 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 为行的 $m\times n$ 矩阵. 这些行向量表示在 n 维欧几里得空间 E_n 中从原点出发的向量. E_n 中由 m 个向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 张成的平行六面体 Π 是由所有形为

$$t_1\boldsymbol{\alpha}_1+\cdots+t_m\boldsymbol{\alpha}_m \quad (0\leqslant t_i\leqslant 1; i=1,\cdots,m)$$

的向量组成. (对 $m=n=3$ 的情形画出图形, 你会得到一个与立方体仿射等价的平行六面体!) 这样就建立起 $m\times n$ 实矩阵与 n 维空间中的 m 维平行六面体之间的对应, $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 称为这个平行六面体 Π 的棱.

这个平行六面体的 m 维体积 $V(\Pi)$(包括当 $m=1$ 时的长度和当 $m=2$ 时的面积这两种特殊情形) 可以对 m 用归纳法来定义. 设以 $\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 为棱的平行六面体称为 Π 的底. $\boldsymbol{\alpha}_1$ 的与 $\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 正交的分量称为高, 它是通过把剩下的一条棱 $\boldsymbol{\alpha}_1$ 写成两个分量 γ 与 $\boldsymbol{\beta}$ 的和来求得, 其中分量 γ 在 $\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 张成的空间 S_{m-1} 中, $\boldsymbol{\beta}$ 与 S_{m-1} 正交 (见图 1, 根据 7.11 节, 这总是可能的):

$$\boldsymbol{\alpha}_1 = \boldsymbol{\beta}+\gamma,\ \boldsymbol{\beta}\perp S_{m-1},\ \gamma 在 S_{m-1}中. \tag{13}$$

Π 的体积就定义为底的 $m-1$ 维体积与高的长度 $|\boldsymbol{\beta}|$ 的乘积.

定理 7　以 $\alpha_1,\cdots,\alpha_m$ 为棱的平行六面体的体积的平方是行列式 $|AA^{\mathrm{T}}|$, 其中 A 是以向量 α_i 的各坐标[①]为第 i 行元素的矩阵.

注　因为 A 的行置换是用 PA 代替 A, 其中 P 是满足 $|P|=|P^{\mathrm{T}}|=\pm1$ 的 $m\times m$ 置换矩阵, 又因为

$$|(PA)(PA)^{\mathrm{T}}|=|P|\cdot|AA^{\mathrm{T}}|\cdot|P^{\mathrm{T}}|=|AA^{\mathrm{T}}|,$$

所以 Π 的 “体积” 与它的底是由哪 $m-1$ 个向量张成的无关.

证明　因为 A 是 $m\times n$ 矩阵, 所以乘积 AA^{T} 是一个 $m\times m$ 方阵. 现在我们对 m 归纳进行论证. 当 $m=1$ 时, 矩阵 A 是行矩阵, “内积”$AA^{\mathrm{T}}=(\alpha_1,\ \alpha_1)$ 是 α_1 的长度的平方, 满足要求. 假定对于 $m-1$ 行的矩阵, 这个定理是正确的, 我们来考虑 m 行的情形. 像 (13) 式那样, 第一行 A_1 可以写成 $A_1=B_1+C_1$, 这里 “高”B_1 同每个行向量 $A_2,\cdots,A_m$ 都正交 (即 $B_1A_i^{\mathrm{T}}=0,i=2,\cdots,m$), 而 $C_2=c_2A_2+\cdots+c_mA_m$ 是 $A_2,\cdots,A_m$ 的线性组合. 从 A 的第一行逐次减去第 i 行的 c_i 倍 $(i=2,\cdots,m)$, 这就把 A 变成新的矩阵 A^*, 它的第一行是 B_1. 而且, 这些初等行运算每一个都相当于用一个行列式为 1 的初等矩阵左乘矩阵 A, 因此 $A^*=PA$, 这里 $|P|=1$, 并且有 $|A^*A^{*\mathrm{T}}|=|PAA^{\mathrm{T}}P^{\mathrm{T}}|=|P||AA^{\mathrm{T}}||P^{\mathrm{T}}|=|AA^{\mathrm{T}}|$. 如果 D 是由 A^* 的 $m-1$ 行 $A_2,\cdots,A_m$ 组成的矩阵块, 那么

$$A^*A^{*\mathrm{T}}=\begin{pmatrix}B_1\\D\end{pmatrix}(B_1^{\mathrm{T}}D^{\mathrm{T}})=\begin{pmatrix}B_1B_1^{\mathrm{T}} & B_1D^{\mathrm{T}}\\ DB_1^{\mathrm{T}} & DD^{\mathrm{T}}\end{pmatrix}=\begin{pmatrix}B_1B_1^{\mathrm{T}} & O\\ O & DD^{\mathrm{T}}\end{pmatrix}$$

这里 $B_1D^{\mathrm{T}}=O$ 是因为对 D 的每行 A_i 有 $B_1A_i^{\mathrm{T}}=0$. 根据 (6) 式, 行列式就是

$$|AA^{\mathrm{T}}|=|A^*A^{*\mathrm{T}}|=(B_1B_1^{\mathrm{T}})|DD^{\mathrm{T}}|.$$

这里 D 是矩阵, 它的所有行 $A_2,\cdots,A_m$ 张成 Π 的底, 所以根据归纳法假设, $|DD^{\mathrm{T}}|$ 是底的体积的平方. 此外, 标量 B_1B^{T} 是高的长度的平方, 所以我们得到所要求的关于 AA^{T} 的底 × 高的公式.　证毕

在行数为 n 的特殊情形中, 显然有 $|AA^{\mathrm{T}}|=|A|\cdot|A^{\mathrm{T}}|=|A|^2$, 于是我们就证明了[②]:

定理 8　设 A 是任意 $n\times n$ 实矩阵, 它的各行为 $\alpha_1,\cdots,\alpha_n$, 则 A 的行列式 (除了可能相差一个正负号外) 是 E_n 中以 $\alpha_1,\cdots,\alpha_n$ 为棱的平行六面体的体积.

① 整个 10.3 节中, 向量的坐标都是对于一组固定的标准正交基来取的. 当 $m>n$, 定理 7 退化为方程 $0=0$.

② 证明这个定理的基本步骤最初是由 J. S. Frame 教授提出的.

行的任意置换并不改变行列式的绝对值, 所以这个定理还表明, 平行六面体的体积定义与棱的排列顺序无关. 当 $m < n$ 时, 这个论证还可应用到定理 7 的公式. 当 $m = n$ 时, 行列式 $|\boldsymbol{A}|$ 常常称为以 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 为棱的平行六面体的"带符号"的体积. 用任意奇置换可使它的符号改变.

定理 9 *一个 n 维欧几里得向量空间的线性变换 $\boldsymbol{Y} = \boldsymbol{X}P$, 使得所有 n 维平行六面体的体积乘上因子 $\pm|\boldsymbol{P}|$.*

证明 考虑一个平行六面体, 它分别以行向量 $\boldsymbol{A}_1, \cdots, \boldsymbol{A}_n$ 为 n 个棱. 行向量 $\boldsymbol{A}_1, \cdots, \boldsymbol{A}_n$ 被变成 $\boldsymbol{A}_1\boldsymbol{P}, \cdots, \boldsymbol{A}_n\boldsymbol{P}$: 含有这些新行向量的矩阵就是矩阵乘积 $\boldsymbol{AP}$, 其中 $\boldsymbol{A}$ 具有行 $\boldsymbol{A}_1, \cdots, \boldsymbol{A}_n$. 那么新的"带符号"的体积等于 $|\boldsymbol{AP}| = |\boldsymbol{A}||\boldsymbol{P}|$, 这里 $|\boldsymbol{A}|$ 是原来平行六面体的体积.

由此推出, 变换 $\boldsymbol{Y} = \boldsymbol{XP}$ 使"带符号"的体积保持不变当且仅当变换矩阵满足 $|\boldsymbol{P}| = 1$. 所有具有这种性质的矩阵 (或所有变换) 组成的集合称为*么模群*. 有时这种群扩大成包含所有满足 $|\boldsymbol{P}| = \pm 1$ 的矩阵 (也就是, 保持体积绝对值不变的所有变换).

n 维欧几里得空间中, 任意区域 f 的体积可以粗略地定义如下: 用一组有限个给定形状和方位的平行六面体 $\Pi_1, \cdots, \Pi_s$ 外接于 f, 然后求和 $\Sigma V(\Pi_i)$, 定义 f 的体积是所有这种平行六面体的不同集合的体积和的最大下界 (第 4 章). (在积分学中, 这一般是可以做到的, 那些平行六面体是其边平行于坐标轴的立方体.)

根据定理 9, 具有矩阵 $\boldsymbol{P}$ 的线性变换以 $1 : |\boldsymbol{P}|$ 的比例改变任意平行六面体的体积, 因此它也以同样的比例改变 f 的体积. 因为平移使体积保持不变, 所以我们得到下面结果.

推论 *仿射变换 $\boldsymbol{Y} = \boldsymbol{XP} + \boldsymbol{K}$ 使所有体积乘上一个因子 $|\boldsymbol{P}|$(或更确切地说, 是乘上 $|\boldsymbol{P}|$ 的绝对值).*

习 题

1. (a) 计算平面上以 (0, 0), (3, 0), (1, 4) 和 (4, 4) 为顶点的平行四边形的面积.
   (b) 计算空间中以 (0, 2, 0), (2, 0, 0), (1, 1, 5) 和 (0, 0, 0) 为相邻顶点的平行六面体的体积.
2. 证明: 任意三角形的三条中线把三角形分成面积相等的六个部分. (*提示*: 用仿射变换把三角形简化为等边三角形的情形.)
3. 证明: 任意平行四边形的两条对角线把平行四边形分为面积相等的四个部分.
4. (a) 设 P 是平行四边形对角线的交点, 证明: 过 P 的任意直线把平行四边形分成面积相等的两个部分.
   (b) 把这个结果推广到三维情形.
5. 描述三个平面, 它们把四面体分为体积相等的六个部分.

6. 用平面三角直接证明: 由向量 $\boldsymbol{\xi}=(x_1,x_2)$ 和 $\boldsymbol{\eta}=(y_1,y_2)$ 张成的平行四边形的面积 A 满足等式

$$A^2=\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}^2.$$

7. (a) 证明: 如果 E_n 中的 m 个向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 线性相关, 那么由这 m 个向量张成的平行六面体的 m 维体积等于零.
   (b) 叙述并证明这个结果的逆命题.
8. 证明: 在正交矩阵群中, 所有满足 $|\boldsymbol{A}|=1$ 的矩阵 (称为 "真" 正交矩阵) 构成指数为 2 的正规子群.
9. (a) 证明: 对应 $\boldsymbol{A}\longmapsto|\boldsymbol{A}|$ 把全线性群同态地映射到由非零标量组成的乘法群.
   (b) 证明: 么模群是全线性群的正规子群.
   (c) 广义么模群 (即所有满足 $|\boldsymbol{P}|=\pm1$ 的矩阵 $\boldsymbol{P}$) 是全线性群的正规子群吗?
10. (a) 证明: 如果 $\boldsymbol{A}$ 是以 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 为行的任意矩阵, 那么 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ 是由内积构成的矩阵 $(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)$.
    (b) 用 (a) 证明: 如果 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 是一组正交向量, 那么

$$|\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}|=(|\boldsymbol{\alpha}_1|\cdots|\boldsymbol{\alpha}_m|)^2.$$

11. (a) 如果 $\boldsymbol{A}$ 是 $m\times n$ 实矩阵, 利用定理 7 的证明来证明 $|\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}|\geqslant 0$. 证明: 当 $m=2$ 时, 这个结果就是 7.10 节定理 18 中的施瓦兹不等式.
    (b) 证明: 以 (0, 0, 0), (x_1,y_1,z_1) 和 (x_2,y_2,z_2) 为顶点的三角形面积是 $\frac{1}{2}|\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}|^{\frac{1}{2}}$ 其中 $\boldsymbol{A}=\begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{pmatrix}$.
    *(c) 以沿 x 轴、y 轴和 z 轴的三个单位线段为棱的四面体的体积为 $\frac{1}{6}$.

    证明: 以 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 为顶点的四面体的体积是 $\frac{1}{6}|\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}}|^{\frac{1}{2}}$, 其中 $\boldsymbol{B}$ 是 $3\times n$ 矩阵, 它的三行分别为 $\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_3-\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_4-\boldsymbol{\alpha}_1$.
    *(d) 推广到高维的 "四面体".

*12. 设 $-K\leqslant a_{ij}\leqslant K(i,j=1,\cdots,n)$.
    (a) 证明: 如果 $\boldsymbol{\alpha}_i=(a_{i1},\cdots,a_{in})$, 那么 $|\boldsymbol{\alpha}_i|\leqslant K\sqrt{n}$.
    (b) 证明: $|\boldsymbol{A}|\leqslant|\boldsymbol{\alpha}_1|\,|\boldsymbol{\alpha}_2|\cdots|\boldsymbol{\alpha}_n|\leqslant K^n n^{\frac{n}{2}}$ (阿达玛 (Hadamard) 行列式定理).

10.4 特征多项式

我们已经看到 (9.2 节定理 5) λ 是 $n\times n$ 矩阵 $\boldsymbol{A}$ 的特征根 (特征值) 当且仅当矩阵 $\boldsymbol{A}-\lambda\boldsymbol{I}$ 是奇异的. 根据定理 3, 也就是当且仅当 $|\boldsymbol{A}-\lambda\boldsymbol{I}|=0$, 这就证明了下面的引理.

引理 矩阵 $\boldsymbol{A}$ 的特征根 (特征值) 是满足 $|\boldsymbol{A}-\lambda\boldsymbol{I}|=0$ 的标量 λ.

在化简矩阵为对角型时, 如果化简是可能的, 这个引理提供了一个直接的方法.

例 设 $\boldsymbol{A}$ 是实对称矩阵

$$\boldsymbol{A}=\begin{pmatrix}1 & 3 & 0\\ 3 & -2 & -1\\ 0 & -1 & 1\end{pmatrix},$$

那么按第一行的子式展开 $|\boldsymbol{A}-\lambda\boldsymbol{I}|$,

$$|\boldsymbol{A}-\lambda\boldsymbol{I}|=\begin{vmatrix}1-\lambda & 3 & 0\\ 3 & -2-\lambda & -1\\ 0 & -1 & 1-\lambda\end{vmatrix}=-\lambda^3+13\lambda-12.$$

进行因式分解得到 $|\boldsymbol{A}-\lambda\boldsymbol{I}|=-(\lambda-1)(\lambda+4)(\lambda-3)$, 所以 $\boldsymbol{A}$ 的特征根是 1, 3, −4. (一般地, 为了求 3×3 矩阵的特征根, 我们必须解一个三次方程, 像 4.4 节或者 5.5 节中那样求解.) 对每个特征根都存在变换 $T_{\boldsymbol{A}}$ 的一个特征向量.

因为

$$(x,y,z)T_{\boldsymbol{A}}=(x+3y,3x-2y-z,-y+z),$$

向量 $\boldsymbol{\xi}=(x,\ y,\ z)$ 是属于特征根 $\lambda=1$ 的特征向量当且仅当 $x+3y=x$, $3x-2y-z=y$, $-y+z=z$; 即当且仅当 $y=0$ 和 $z=3x$, 于是得到 $\boldsymbol{\xi}=(x,0,3x)$. 类似地, $\boldsymbol{\xi}=(x,y,z)$ 是属于特征根 $\lambda=3$ 的特征向量当且仅当 $x+3y=3x$, $3x-2y-z=3y$, $-y+z=3z$, 于是 $\boldsymbol{\xi}$ 只能是向量 (3, 2, −1) 的标量倍. 同样, 属于特征根 $\lambda=-4$ 的特征向量是向量 (−3, 5, 1) 的标量倍. 这三个特征向量

$$(1,0,3),\ (3,2,-1),\ (-3,5,1)$$

相互正交, 因此它们线性无关. 以这三个向量为行的矩阵 $\boldsymbol{P}$ 是非奇异的. 对于由这三个向量构成的一组新基, 变换 $T_{\boldsymbol{A}}$ 是非奇异的, 它对应的矩阵是 $\boldsymbol{PAP}^{-1}$(参看 9.2 节定理 3′ 和定理 4). 我们还可以把这组基标准化, 得到特征向量的标准正交基

$$\alpha_1=\frac{1}{\sqrt{10}}(1,0,3),\quad \alpha_2=\frac{1}{14}(3,2,-1),\quad \alpha_3=\frac{1}{\sqrt{35}}(-3,5,1),$$

以这三个向量为行的矩阵 $\boldsymbol{Q}$ 是正交的, 并且 $\boldsymbol{QAQ}^{-1}=\boldsymbol{QAQ}^{\mathrm{T}}$ 是对角矩阵, 其对角线元素为 1, 3, −4.

上述 3×3 对称矩阵 $\boldsymbol{A}$ 是二次型 $x^2+6xy-2y^2-2yz+z^2$ 的矩阵. 前面的分析表明, 这个二次型对于标准正交基 ("主轴")$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$, 表为对角型 $x^2+3y^2-4z^2$.

一般地, 设 $\boldsymbol{A}$ 是任意 $n\times n$ 矩阵. 因为行列式是一个多项式, 对每一行的元素来说是线性的, 所以行列式 $|\boldsymbol{A}-\lambda\boldsymbol{I}|$ 是未定元 λ 的 n 次多项式:

$$|\boldsymbol{A}-\lambda\boldsymbol{I}|=(-1)^n\lambda^n+b_{n-1}\lambda^{n-1}+\cdots+b_1\lambda+b_0. \tag{14}$$

我们将定义 $\boldsymbol{A}$ 的特征多项式是 $c_{\boldsymbol{A}}(\lambda)=|\boldsymbol{A}-\lambda\boldsymbol{I}|$, $\boldsymbol{A}$ 的特征方程是方程 $|\boldsymbol{A}-\lambda\boldsymbol{I}|=0$. 现在将上面的引理重述如下：

定理 10 矩阵 $\boldsymbol{A}$ 的特征根(特征值)是 $\boldsymbol{A}$ 的特征方程的根.

因为任意复多项式至少有一个根, 所以我们得出下面推论：

推论 在复数域上, 一个线性变换至少有一个 (非零) 特征向量.

定理 11 相似矩阵有相同的特征多项式.

证明 设两个相似的矩阵是 $\boldsymbol{A}$ 和 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$. 因为 $|\boldsymbol{P}^{-1}|=|\boldsymbol{P}|^{-1}$, 并且 $|\boldsymbol{P}|$ 是标量, 所以可以变换, 于是由行列式乘法法则得出

$$\begin{aligned}|\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}-\lambda\boldsymbol{I}|&=|\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}-\lambda\boldsymbol{P}^{-1}\boldsymbol{I}\boldsymbol{P}|=|\boldsymbol{P}^{-1}(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{P}|\\&=|\boldsymbol{P}^{-1}|\cdot|\boldsymbol{A}-\lambda\boldsymbol{I}|\cdot|\boldsymbol{P}|=|\boldsymbol{A}-\lambda\boldsymbol{I}|.\end{aligned}$$

因此我们得到一个推论：$|\boldsymbol{A}-\lambda\boldsymbol{I}|$ 的逐个系数

$$\begin{aligned}&b_0=|\boldsymbol{A}|,\\&b_1=\cdots\\&\cdots\cdots\\&b_{n-2}=(-1)^n\sum_{i<j}(a_{ii}a_{jj}-a_{ij}a_{ji}),\\&b_{n-1}=(-1)^n(a_{11}+\cdots+a_{nn})\end{aligned}$$

是矩阵 $\boldsymbol{A}$ 在相似群 $\boldsymbol{A}\longmapsto\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 之下的不变量. 关于 b_i 的某些适当的多项式给出另外一些有用的不变量. 在这些不变量中, 有一个是

$$\sum_{i,j=1}^{n}a_{ij}a_{ji}=\sum_{i=1}^{n}a_{ii}^2+2\sum_{i<j}a_{ij}a_{ji}=b_{n-1}^2+(-1)^{n-1}2b_{n-2}.$$

对于对称矩阵的情形, 这个不变量就是 Σa_{ij}^2.

因为 $|\boldsymbol{A}^{\mathrm{T}}-\lambda\boldsymbol{I}^{\mathrm{T}}|=|(\boldsymbol{A}-\lambda\boldsymbol{I})^{\mathrm{T}}|=|(\boldsymbol{A}-\lambda\boldsymbol{I})|$(根据定理 1), 我们还有

推论 矩阵 $\boldsymbol{A}$ 和它的转置矩阵 $\boldsymbol{A}^{\mathrm{T}}$ 具有相同的特征多项式, 因此具有相同的特征根.

定理 12 对角线元素为 $d_1,d_2,\cdots,d_n$ 的三角形矩阵 $\boldsymbol{T}$ 的特征多项式是

$$|\boldsymbol{T}-\lambda\boldsymbol{I}|=(d_1-\lambda)(d_2-\lambda)\cdots(d_n-\lambda).$$

因为 $\boldsymbol{T}-\lambda\boldsymbol{I}$ 本身也是三角形矩阵, 所以根据 10.1 节的引理 2 便得到定理的证明. 由此得到推论：对角线元素 (可以重复出现) 集合是由特征多项式的全体根 (有的是重根) 组成. 因此对于两个相似的对角矩阵, 对角线元素集合和每个对角线元素出现的次数都是一样的. 这可以叙述如下：

推论 两个对角矩阵是相似的当且仅当只是它们的对角线元素的次序不同.

相似性的这个性质给实二次型的正交变换 (9.10 节) 以新的解释. 如果具有矩阵 $\boldsymbol{A}$ 的二次型 $\boldsymbol{XAX}^{\mathrm{T}}$ 通过正交变换 $\boldsymbol{Z}=\boldsymbol{XP}$ 化成对角型 $\lambda_1 z_1^2+\cdots+\lambda_n z_n^2$, 那么这个新二次型的对角矩阵是 $\boldsymbol{D}=\boldsymbol{PAP}^{\mathrm{T}}$. 因为 $\boldsymbol{P}$ 是正交矩阵, 所以 $\boldsymbol{P}^{\mathrm{T}}=\boldsymbol{P}^{-1}$, $\boldsymbol{D}=\boldsymbol{PAP}^{-1}$, 因此新矩阵 $\boldsymbol{D}$ 与原矩阵 $\boldsymbol{A}$ 是相似的, 所以 $\boldsymbol{D}$ 的特征值 $\lambda_1,\cdots,\lambda_n$ 与已知矩阵 $\boldsymbol{A}$ 的特征值相同. 这就给出下面 9.10 节定理 21 的更强的形式:

定理 13 任意实二次型 $\boldsymbol{XAX}^{\mathrm{T}}$ 可以用正交变换化成对角型 $\lambda_1 z_1^2+\cdots+\lambda_n z_n^2$, 其中系数 λ_i 是 $\boldsymbol{A}$ 的特征方程 $|\boldsymbol{A}-\lambda\boldsymbol{I}|=(\lambda_1-\lambda)\cdots(\lambda_n-\lambda)=0$ 的根.

特征方程以及它的根是由矩阵 $\boldsymbol{A}$ 唯一确定的. 这就证明了对角型本质上的唯一性, 并且给出直接计算系数的方法. 知道了系数以后, 我们还可以按照上面指出的方法计算出与此相联系的特征向量作为主轴.

因为我们知道, 任意实对称矩阵正交等价于实对角矩阵, 所以我们得到

推论 实对称矩阵的所有特征值都是实的.

附注 如果 $\boldsymbol{A}$ 是对称矩阵, 那么属于不同特征值 $\lambda_1\neq\lambda_2$ 的特征向量 $\boldsymbol{X}_1$ 和 $\boldsymbol{X}_2$ 必然是正交的, 因为双线性表达式 $\boldsymbol{X}_1\boldsymbol{AX}_2^{\mathrm{T}}$ 可以按两种方法来计算:

$$(\boldsymbol{X}_1\boldsymbol{A})\boldsymbol{X}_2^{\mathrm{T}}=\lambda_1(\boldsymbol{X}_1\boldsymbol{X}_2^{\mathrm{T}}),\qquad \boldsymbol{X}_1(\boldsymbol{AX}_2^{\mathrm{T}})=\boldsymbol{X}_1(\boldsymbol{X}_2\boldsymbol{A})^{\mathrm{T}}=\lambda_2(\boldsymbol{X}_1\boldsymbol{X}_2^{\mathrm{T}}).$$

因为 $\lambda_1\neq\lambda_2$, 所以 $\boldsymbol{X}_1\boldsymbol{X}_2^{\mathrm{T}}$ 一定是零, 因此 $\boldsymbol{X}_1$ 与 $\boldsymbol{X}_2$ 正交.

因此, 如果 $n\times n$ 对称矩阵 $\boldsymbol{A}$ 有 n 个不同的特征值 $\lambda_1,\cdots,\lambda_n$, 那么任意 n 个相应的特征向量 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n$ 是正交的, 并且以单位向量 $\dfrac{\boldsymbol{X}_1}{|\boldsymbol{X}_1|},\cdots,\dfrac{\boldsymbol{X}_n}{|\boldsymbol{X}_n|}$ 为行的正交矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{PAT}^{\mathrm{T}}=\boldsymbol{PAP}^{-1}$ 成为对角矩阵.

习 题

1. 设 $\boldsymbol{D}$ 是对角矩阵, 其对角线元素为 3, 1, −1, 而 $\boldsymbol{P}$ 是以 (1, 2, −3), (0, −1, 4), (0, 0, 1) 为行的三角形矩阵. 计算 $\boldsymbol{P}^{-1}\boldsymbol{DP}$ 的特征方程, 并与 $\boldsymbol{D}$ 的特征方程进行比较.
2. 计算下列各矩阵的特征值和特征向量

$$\text{(a)}\begin{pmatrix}-1&2&2\\2&2&2\\-3&-6&-6\end{pmatrix},\quad \text{(b)}\begin{pmatrix}3&2&2\\1&4&1\\-2&-4&-1\end{pmatrix},\quad \text{(c)}\begin{pmatrix}4&9&0\\0&-2&8\\0&0&7\end{pmatrix}.$$

3. 求二次型 $xy+yz+zx+x+y+z=1$ 的主轴长度.
4. 写出在正交变换之下等价于下面给出的表达式的对角二次型:
   (a) $-2x^2-11y^2-5z^2+4xy+16yz+20xz$. (提示: 证明所有整数特征值是 9 的倍数.)
   (b) $3x^2-y^2-3z^2-t^2-4xz-10yt$.
5. 写出把习题 4 的每个二次型化为等价的对角型的正交变换.
6. 求出 2 × 2 矩阵的特征值相等的充分必要条件.

7. 求出其特征值为 1 和 -1 的全部 2×2 矩阵.
8. 证明：如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是方阵, 那么矩阵 $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{C} & \boldsymbol{B} \end{pmatrix}$ 的特征多项式是 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的特征多项式的乘积.
9. 证明：(14) 式中的 $b_{n-1} = \pm(a_{11} + \cdots + a_{nn})$, (不变量 $a_{11} + \cdots + a_{nn}$ 称为 $\boldsymbol{A}$ 的迹.)
10. 证明：(14) 式中的 $b_{n-2} = (-1)^n \sum_{i<j} (a_{ii}a_{jj} - a_{ij}a_{ji})$.
11. 证明：对于对称矩阵有 $\Sigma a_{ij}^2 = b_{n-1}^2 + (-1)^{n-1} 2b_{n-2}$.
12. 从定义直接证明：实对称矩阵 $\boldsymbol{A}$ 的所有特征值都是实的. (提示：对于特征向量 $\boldsymbol{X}$, 证明：$\boldsymbol{X}\boldsymbol{A}(\boldsymbol{X}^*)^{\mathrm{T}} = \lambda \boldsymbol{X}(\boldsymbol{X}^*)^{\mathrm{T}}) = \lambda^* \boldsymbol{X}(\boldsymbol{X}^{\mathrm{T}})^*$, 这里 $\boldsymbol{X}^*$ 表示 $\boldsymbol{X}$ 的复共轭.
13. (a) 证明：埃尔米特矩阵的所有特征值都是实的.
    (b) 证明：全体特征向量张成所有向量组成的空间.

*14. 证明：每个酉矩阵 $\boldsymbol{U}$ 有一个满足 $\boldsymbol{\xi U} = d\boldsymbol{\xi}$ 的特征向量 $\boldsymbol{\xi}$, 其中 $|d| = 1$.

15. (a) 证明：如果矩阵 $\boldsymbol{A}$ 对于特征值 λ_j 有 r 个线性无关的特征向量, 那么特征多项式 $c_{\boldsymbol{A}}(\lambda)$ 是 $(\lambda - \lambda_j)^r$ 的一个倍式.
    (b) 对任意 r, 构造一个 $r \times r$ 矩阵 $\boldsymbol{A}$, 使得 $c_{\boldsymbol{A}}(\lambda) = (\lambda - \lambda_1)^r$, 而对于特征值 λ_1, 没有两个线性无关的特征向量.
16. 通过对线性变换 $\boldsymbol{X} \longmapsto \boldsymbol{X}\boldsymbol{A}$ 所进行的如下分析, 证明实对称矩阵 $\boldsymbol{A}$ 的主轴定理.
    (a) 矩阵 $\boldsymbol{A}$ 有一个长度为 1 的特征向量 $\boldsymbol{\alpha}_1$.
    (b) 如果把 $\boldsymbol{\alpha}_1$ 选作新的标准正交基的第一个向量, 那么已知变换的新矩阵的第一行和第一列除了第一个元素外其他元素都是零.
    (c) 用归纳法继续完成上面的论证.

*17. 证明：椭球 $\Sigma a_{ij}x_ix_j \leqslant 1$ 的体积等于 $\frac{4\pi}{3}|\boldsymbol{A}|^{-\frac{1}{2}}$, 其中 $\boldsymbol{A} = (a_{ij})$. (提示：变换到主轴上, 并用定理 9.)

10.5 极小多项式

矩阵在相似变换下标准型的构造依赖于对矩阵或相应的变换所满足的多项式方程的研究. 尤其是, 设 V 是域 F 上的 n 维向量空间, $T: V \to V$ 是 V 的线性变换, T 的各次幂 T^m 也是 V 的线性变换. 因为线性变换还可以相加或乘上一个标量, 所以对每个形为 $f(x) = a_0 + a_1x + \cdots + a_kx^k$(其中系数 $a_i \in F$) 的多项式, 我们可以考虑相应的关于 T 的多项式

$$f(T) = a_0I + a_1T + \cdots + a_kT^k. \tag{15}$$

它表示一个线性变换 $f(T): V \to V$, 特别是常数多项式 $f(x) = 1$ 产生恒等变换 $I: V \to V$. 因为 T 的幂是可交换的 $(T^mT^q = T^qT^m = T^{m+q})$, 所以多项式 $f(T)$ 可以像多项式 $f(x)$ 那样相加和相乘.

类似地, 元素在 F 中的每个 $n\times n$ 矩阵 $\boldsymbol{A}$ 产生出 $\boldsymbol{A}$ 的多项式

$$f(\boldsymbol{A})=a_0\boldsymbol{I}+a_1\boldsymbol{A}+\cdots+a_k\boldsymbol{A}^k, \tag{16}$$

它们还是元素在 F 中的 $n\times n$ 矩阵. 因为恰好存在 n^2 个线性无关的 F 上的 $n\times n$ 矩阵, 所以 n^2+1 个矩阵 $\boldsymbol{I},\boldsymbol{A},\cdots,\boldsymbol{A}^{n^2}$ 一定线性相关, 并且这种相关关系提供了一个次数至多为 n^2 的非零多项式 $f(x)$, 满足 $f(\boldsymbol{A})=\boldsymbol{O}$. 由于 $n\times n$ 矩阵和 V_n 的线性变换之间存在同构 $\boldsymbol{A}\longmapsto T_{\boldsymbol{A}}$, 所以对 n 维向量空间 V 的每个线性变换 T 也存在一个非零多项式 $f(x)$ 使得 $f(T)=\boldsymbol{O}$.

定理 14 对域 F 上有限维向量空间 V 的每个线性变换 T, 使得 $f(T)=\boldsymbol{O}$ 的 F 上的多项式 $f(x)$ 是唯一的首一多项式 $m(x)$ 的倍式.

证明 考虑 F 上满足 $f(T)=O$ 的所有多项式 $f(x)$ 组成的集合 M. 我们刚刚看到 M 包含非零多项式, 而且 M 关于加法、减法和用任意多项式 $g(x)$ 相乘的三种运算是封闭的, 因此 M 是环 $F[x]$ 中的一个理想. 因此根据 3.8 节定理 11, M 是由满足 $m(T)=O$ 的次数最小的首一多项式 $m(x)$ 的全体倍式组成.

我们称 $m(x)$ 是 T 的极小多项式. 它是具有下述性质的首一多项式:

$$m(T)=\boldsymbol{O};\qquad \text{由}f(T)=\boldsymbol{O}\ \text{推出}m(x)|f(x), \tag{17}$$

其中记号 $m(x)|f(x)$ 表示在多项式环 $F[x]$ 中 $m(x)$ 整除 $f(x)$(像第 3 章那样). 对 $n\times n$ 矩阵 $\boldsymbol{A}$ 的极小多项式可做类似的描述, 它等同于 F^n 上相应变换 $T_{\boldsymbol{A}}$ 的极小多项式. 因为相似的矩阵是同一个线性变换的不同表示, 所以我们有

推论 域 F 上的相似矩阵有相同的极小多项式.

作为一个例子, 我们考虑幂零变换 (或幂零矩阵), 即对某个 m 满足 $T^m=O$ 的线性变换 T. 因为 T 满足 $T^m=O$, 所以它的极小多项式是 x^h, 其中 h 是某个整数. 事实上 h 是满足 $T^h=O$ 的最小正整数.

特殊情形是假定 $h=n$. 因为 $T^{h-1}=T^{n-1}\neq O$, 所以存在一个向量 $\boldsymbol{\alpha}$ 满足 $\boldsymbol{\alpha}T^{n-1}\neq\boldsymbol{0}$. 我们断言, $\boldsymbol{\alpha},\boldsymbol{\alpha}T,\boldsymbol{\alpha}T^2,\cdots,\boldsymbol{\alpha}T^{n-1}$ 这 n 个向量是线性无关的. 如果不然, 则存在线性相关关系 $\boldsymbol{0}=a_0\boldsymbol{\alpha}+a_1\boldsymbol{\alpha}T+\cdots+a_{n-1}\boldsymbol{\alpha}T^{n-1}$, 其系数不全为零. 如果 a_j 是第一个不为零的系数, 那么我们就用 T^{n-j-1} 乘这个方程两边得到

$$\boldsymbol{0}=\boldsymbol{0}T^{n-j-1}=a_j\boldsymbol{\alpha}T^jT^{n-j-1}=a_j\boldsymbol{\alpha}T^{n-1},$$

而这里选取的 $\boldsymbol{\alpha}$ 满足 $\boldsymbol{\alpha}T^{n-1}\neq\boldsymbol{0}$, 因此 $a_j=\boldsymbol{0}$, 矛盾. 当选取 $\boldsymbol{\alpha},\boldsymbol{\alpha}T,\cdots,\boldsymbol{\alpha}T^{h-1}$ 这些线性无关向量作为基时, T 把每个基向量变到下一个基向量, 并把最后一个基向

量变为零向量，因此 T 用 $n\times n$ 矩阵表示就是

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

这个矩阵中非零元素都是 1，它们分布在紧挨主对角线上面的一条对角线上. 这个矩阵显然是幂零矩阵，它称为多项式 x^n 的“友矩阵”.

更一般地，对每个 n 次首一多项式

$$g(x)=c_0+c_1+\cdots+c_{n-1}x^{n-1}+x^n,$$

我们可以构造一个以 $g(x)$ 为极小多项式的 $n\times n$ 矩阵. 这个矩阵称为 $g(x)$ 的友矩阵，例如 $n=4$ 时，

$$\boldsymbol{C}_g=\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -c_0 & -c_1 & -c_2 & -c_3 \end{pmatrix}; \tag{18}$$

对于任意 $n,\boldsymbol{C}_g$ 的紧挨主对角线上面的一条对角线上的元素都是 1，最后一行的元素是 $-c_0,\cdots,-c_{n-1}$，其余元素都是零.

定理 15　*对每个首一多项多 $g(x)$，友矩阵 $\boldsymbol{C}_g$ 有极小多项式 $g(x)$ 和特征多项式 $(-1)^n g(\lambda)$.*

证明　设 T 是 F^n 的用形如 (18) 的友矩阵 $\boldsymbol{C}_g$ 所表示的线性变换. 因为这个矩阵的各行分别是 F^n 的单位向量 $\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n$ 的变换式的坐标，所以我们有

$$\boldsymbol{\varepsilon}_1 T=\boldsymbol{\varepsilon}_2,\ \cdots,\ \boldsymbol{\varepsilon}_{n-1}T=\boldsymbol{\varepsilon}_n,\quad \boldsymbol{\varepsilon}_n T=-c_0\boldsymbol{\varepsilon}_1-\cdots-c_{n-1}\boldsymbol{\varepsilon}_n.$$

换句话说，向量 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_1 T,\cdots,\boldsymbol{\varepsilon}_1 T^{n-1}$ 是 F^n 的一组基，所以任意向量 $\boldsymbol{\xi}$ 可以唯一地写成

$$\boldsymbol{\xi}=a_0\boldsymbol{\varepsilon}_1+a_1\boldsymbol{\varepsilon}_1 T+\cdots+a_{n-1}\boldsymbol{\varepsilon}_1 T^{n-1}=\boldsymbol{\varepsilon}_1 f(T), \tag{19}$$

其中 $f(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}$ 是次数至多为 $n-1$ 的多项式. 进一步有 $\boldsymbol{\varepsilon}_1 T^n=-c_0\boldsymbol{\varepsilon}_1-\cdots-c_{n-1}\boldsymbol{\varepsilon}_1 T^{n-1}$，所以 $\boldsymbol{\varepsilon}_1 g(T)=\mathbf{0}$. 因此对任意向量 $\boldsymbol{\xi}$，

$$\boldsymbol{\xi}g(T)=\boldsymbol{\varepsilon}_1 f(T)g(T)=\boldsymbol{\varepsilon}_1 g(T)f(T)=0,$$

这就断言, T 满足首一多项式方程 $g(T)=O$. 对于任意次数较低的多项式 $f(x)\neq O$, 由 (19) 式可知 $\varepsilon_1 f(T)=\boldsymbol{\xi}\neq O$, 因此 $f(T)\neq O$. 于是 $g(x)$ 确实是 $\boldsymbol{C}_g$ 的极小多项式.

$\boldsymbol{C}_g$ 的特征多项式是通过把行列式 $|\boldsymbol{C}_g-\lambda\boldsymbol{I}|$ 按最后一行的子式展开求得的. 因为 $-c_k$ 的子式是三角形矩阵, 它的对角线元素有 k 个是 $-\lambda$, 其他都是 1, 所以 $|\boldsymbol{C}_g-\lambda\boldsymbol{I}|$ 确实等于 $(-1)^n g(\lambda)$, 这里出现符号 $(-1)^n$ 是因为任意 $n\times n$ 矩阵的特征多项式的首项都是 $(-1)^n\lambda^n$.

习　题

1. (a) 证明：任意满足 $\boldsymbol{X}^2=\boldsymbol{O}$ 的 2×2 矩阵 $\boldsymbol{X}$ 相似于矩阵 $\begin{pmatrix}0&1\\0&0\end{pmatrix}$ 或者 $\boldsymbol{X}$ 是 $\begin{pmatrix}0&0\\0&0\end{pmatrix}$.

   (b) 对 3×3 矩阵证明相应的结果.

2. 证明：每个实 2×2 矩阵, 如果它的行列式是负的, 那么它同对角矩阵相似. 给出几何解释.

3. (a) 对任意 $n\times n$ 非奇异矩阵 $\boldsymbol{P}$, 证明：对应 $\boldsymbol{A}\longmapsto\boldsymbol{PAP}^{-1}$ 是所有 $n\times n$ 矩阵 $\boldsymbol{A}$ 的代数的自同构.

   (b) 根据 (a) 直接证明：相似矩阵有相同的极小多项式.

4. 证明：任意对角矩阵的特征多项式是它的极小多项式的倍式. 什么时候两者相同?

*5. (a) 证明：每个行列式为负的 2×2 实正交矩阵表示一个刚体反射. (提示：见习题 2 或 9.4 节)

   (b) 证明：每个行列式为正的 2×2 正交矩阵表示一个刚体旋转.

*6. (a) 证明：任意 3×3 实矩阵 $\boldsymbol{A}$ 有一个实特征向量.

   (b) 证明：任意 3×3 正交矩阵, 在基的正交变换下, 它相似于形为 $\begin{pmatrix}\pm1&0\\0&\boldsymbol{B}\end{pmatrix}$ 的矩阵, 这里 $\boldsymbol{B}$ 是 2×2 正交矩阵.

   (c) 用习题 5 证明：如果 $\boldsymbol{A}$ 是 3×3 正交矩阵, 并且 $|\boldsymbol{A}|>0$, 那么 $\boldsymbol{A}$ 有一个特征值是 1, 而且 $\boldsymbol{A}$ 表示刚体旋转. (这就是欧拉定理.)

*7. 证明：如果 λ 是 $\boldsymbol{A}$ 的特征值, $q(x)$ 是任意多项式, 那么 $q(\lambda)$ 是 $q(\boldsymbol{A})$ 的特征值.

*8. (a) 证明：矩阵

$$\boldsymbol{C}=\begin{pmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\1&0&0&0\end{pmatrix}$$

   的特征值为 ±1, $\pm\mathrm{i}$, 即复四次单位根.

   (b) $\boldsymbol{C}$ 的复特征向量是什么?

   (c) $\boldsymbol{C}$ 与什么样的复对角矩阵相似?

*9. 一个 $n\times n$ 矩阵, 如果它满足对一切 i,j, $a_{ij}=a_{i+1,j+1}$(其中下标 i, j 都取模 n 整数),

则称这个矩阵为轮换矩阵. 证明: 任意轮换矩阵的特征值 $\lambda_1, \cdots, \lambda_n$ 是

$$\lambda_p = a_{11} + a_{12}\omega^p + \cdots + a_{1n}\omega^{(n-1)p},$$

式中 ω 是 n 次本原单位根. (提示: 利用习题 7 和习题 8.)

10.6 凯莱－哈密顿定理

我们现在将要证明, 每个方阵 $\boldsymbol{A}$ 都满足它的特征方程, 也就是说, $\boldsymbol{A}$ 的极小多项式整除 $\boldsymbol{A}$ 的特征多项式.

用矩阵多项式或者 λ 矩阵的概念很容易证明上述性质. λ 矩阵 $\boldsymbol{A}-\lambda\boldsymbol{I}$ 是指以 λ 多项式为元素的矩阵. 把 λ 的同次幂的项合并, 我们可以把任意非零 λ 矩阵 $\boldsymbol{B}(\lambda)$ 写成下面形式

$$\boldsymbol{B}(\lambda) = \boldsymbol{B}_0 + \lambda\boldsymbol{B}_1 + \cdots + \lambda^r\boldsymbol{B}_r,$$

其中 $\boldsymbol{B}_i$ 是常数矩阵, $\boldsymbol{B}_r \neq \boldsymbol{O}$. (等式意味着两边的矩阵各个元素中 λ 的每个相应系数都相等.)

引理 *如果 $\boldsymbol{C} = \boldsymbol{B}(\lambda)(\boldsymbol{A} - \lambda\boldsymbol{I})$ 是常数矩阵, 那么 $\boldsymbol{C} = \boldsymbol{O}$.*

证明 展开 $\boldsymbol{B}(\lambda)(\boldsymbol{A} - \lambda\boldsymbol{I})$, 我们得到

$$-\lambda^{r+1}\boldsymbol{B}_r + \sum_{k=1}^{r}\lambda^k(\boldsymbol{B}_k\boldsymbol{A} - \boldsymbol{B}_{k-1}) + \boldsymbol{B}_0\boldsymbol{A},$$

如果它是常数矩阵, 意味着 $\boldsymbol{B}_r = \boldsymbol{O}, \boldsymbol{B}_k\boldsymbol{A} - \boldsymbol{B}_{k-1} = \boldsymbol{O}, k = 1, \cdots, r$, 故得 $\boldsymbol{B}(\lambda) = \boldsymbol{O}$. 现在结论显然成立.

定理 16 (凯莱–哈密顿) *每个方阵满足它的特征方程.*

这意味着, (14) 的特征多项式 $f(\lambda) = |\boldsymbol{A} - \lambda\boldsymbol{I}|$ 中, 如果 λ 的每个幂 λ^i 用矩阵 $\boldsymbol{A}$ 的同次幂 $\boldsymbol{A}^i$ 代替, (并且 λ^0 用 $\boldsymbol{A}^0 = \boldsymbol{I}$ 代替), 其结果是零:

$$b_0\boldsymbol{I} + b_1\boldsymbol{A} + \cdots + b_{n-1}\boldsymbol{A}^{n-1} + (-1)^n\boldsymbol{A}^n = \boldsymbol{O}. \tag{20}$$

证明 在矩阵 $\boldsymbol{A} - \lambda\boldsymbol{I}$ 中, 每个元素都是 λ 的线性多项式, 所以它的非零子式也是 λ 的 $n-1$ 次或小于 $n-1$ 次的多项式. $\boldsymbol{A} - \lambda\boldsymbol{I}$ 的伴随矩阵 $\boldsymbol{C}$ 的每个元素是这样的子式, 所以这个伴随矩阵可以写成 n 个矩阵的和, 这 n 个矩阵分别包含 λ 的固定的幂 $\lambda^0, \lambda^1, \cdots, \lambda^{n-1}$. 换句话说, 伴随矩阵 $\boldsymbol{C} = \boldsymbol{C}(\lambda)$ 是 λ 矩阵 $\boldsymbol{C} = \boldsymbol{C}(\lambda) = \Sigma\lambda^i\boldsymbol{C}_i$. 根据 (10) 式, $\boldsymbol{A} - \lambda\boldsymbol{I}$ 与它的伴随矩阵的乘积是

$$\boldsymbol{C}(\lambda)(\boldsymbol{A} - \lambda\boldsymbol{I}) = |\boldsymbol{A} - \lambda\boldsymbol{I}|\boldsymbol{I} = f(\lambda)\boldsymbol{I}, \tag{21}$$

这里 $f(\lambda)$ 是特征多项式.

现在注意, 由熟悉的因式分解公式

$$\boldsymbol{A}^i-\lambda^i\boldsymbol{I}=(\boldsymbol{A}^{i-1}+\lambda\boldsymbol{A}^{i-2}+\cdots+\lambda^{i-1}\boldsymbol{I})(\boldsymbol{A}-\lambda\boldsymbol{I}),$$

再利用特征多项式 (14) 的系数 b_i, 我们得到

$$\begin{aligned}f(\boldsymbol{A})-f(\lambda)\boldsymbol{I}&=\sum_{i=0}^{n}b_i\boldsymbol{A}^i-\sum_{i=0}^{n}b_i\lambda^i\boldsymbol{I}=\sum_{i=1}^{n}b_i(\boldsymbol{A}^i-\lambda^i\boldsymbol{I})\\&=\sum_{i=1}^{n}b_i(\boldsymbol{A}^{i-1}+\lambda\boldsymbol{A}^{i-2}+\cdots+\lambda^{i-1}\boldsymbol{I})(\boldsymbol{A}-\lambda\boldsymbol{I}),\end{aligned}$$

其中 $f(\boldsymbol{A})$ 是特征多项式 $f(\lambda)$ 中的 λ 换成 $\boldsymbol{A}$ 而得到的. 上式即

$$f(\boldsymbol{A})-f(\lambda)\boldsymbol{I}=-\boldsymbol{G}(\lambda)(\boldsymbol{A}-\lambda\boldsymbol{I}),\tag{22}$$

其中 $\boldsymbol{G}(\lambda)$ 是一个新的 λ 矩阵. 如果我们把 (22) 与 (21) 相加, 便得

$$[\boldsymbol{C}(\lambda)-\boldsymbol{G}(\lambda)](\boldsymbol{A}-\lambda\boldsymbol{I})=f(\boldsymbol{A}),$$

其中 $f(\boldsymbol{A})$ 是常数矩阵. 根据引理, 这就推出 $f(\boldsymbol{A})=\boldsymbol{O}$.

习　题

1. 通过直接代入证明：每个 2×2 矩阵满足它的特征方程.
2. 证明：如果 $\boldsymbol{A}$ 是非奇异矩阵, 并有特征多项式 (14), 那么 $\boldsymbol{A}$ 的伴随矩阵由

$$-[b_1\boldsymbol{I}+b_2\boldsymbol{A}+\cdots+b_{n-1}\boldsymbol{A}^{n-2}+(-1)^n\boldsymbol{A}^{n-1}]$$

给出.

3. 根据习题 2 的表示法证明：$\boldsymbol{A}^{-1}$ 的特征多项式为

$$(-1)^n\left[\lambda^n+\frac{b_1}{|\boldsymbol{A}|}\lambda^{n-1}+\cdots+\frac{(-1)^n}{|\boldsymbol{A}|}\right].$$

4. (a) 通过直接计算证明：关于严格三角形矩阵的凯莱–哈密顿定理.

   *(b) 对三角形矩阵回答同样问题.
5. (a) 用直接计算证明：(18) 式的 4×4 友矩阵 $\boldsymbol{C}_g$ 满足它的特征方程.

   (b) 对 n 次多项式的友矩阵做同样证明.

10.7　不变子空间与可约性

如果一个线性变换 T 满足一个可以因式分解的多项式方程, 那么表示 T 的矩阵常常能够相应地简化. 例如, 我们假定 T 满足 $T^2=I$(周期是 2), 在基域 F 中

1+1≠0, 所以 $(T-I)(T+I)=O$ 的两个因子是互素的. T 的特征向量包含 $T+I$ 的值域中的全部非零向量 $\boldsymbol{\eta}=\boldsymbol{\xi}(T+I)$, 这因为

$$(\boldsymbol{\xi}(T+I))T=\boldsymbol{\xi}(T^2+T)=\boldsymbol{\xi}(T+I).$$

这些向量 $\boldsymbol{\eta}$ 是属于特征值 1 的. 同样地, $T-I$ 的值域中的所有非零向量是属于特征值 -1 的特征向量, 这因为

$$(\boldsymbol{\xi}(T-1))T=\boldsymbol{\xi}(T^2-T)=\boldsymbol{\xi}(I-T)=-\boldsymbol{\xi}(T-I).$$

而因为 1+1≠ 0, 所以任意向量 $\boldsymbol{\xi}$ 可以写成两个向量的和

$$\boldsymbol{\xi}=\frac{1}{2}[\boldsymbol{\xi}(T+I)-\boldsymbol{\xi}(T-I)].$$

因此属于特征值 ±1 的特征向量张成这个空间, 于是根据 9.2 节定理 4, T 可以用对角线元素是 ±1 的对角矩阵表示.

特别地, 如果对角线元素都是 1, 则 T 是恒等变换, 并且 T 的极小多项式是 $x-1$; 如果对角线元素都是 -1, 则 T 的极小多项式是 $x+1$; 如果对角线元素既有 1 又有 -1, 则 T 的极小多项式是 x^2-1. 这个分析是下面定理的特殊情形.

定理 17 *如果线性变换 $T:V\to V$ 的极小多项式 $m(x)$ 在 V 的基域 F 上可以分解因式为 $m(x)=f(x)g(x)$, 其中 $f(x)$ 和 $g(x)$ 是首一多项式, 并且互素, 那么 V 中任意向量可以唯一地表示成和*

$$\boldsymbol{\xi}=\boldsymbol{\eta}+\boldsymbol{\zeta},\qquad \boldsymbol{\eta}f(T)=\mathbf{0},\qquad \boldsymbol{\zeta}g(T)=\mathbf{0}. \tag{23}$$

证明 因为 f 和 g 是互素的, 所以由欧几里得算法给出系数在 F 中的两个多项式 $h(x)$ 和 $k(x)$, 使得

$$1=h(x)f(x)+k(x)g(x). \tag{24}$$

将 T 代入 x, 得到 $I=h(T)f(T)+k(T)g(T)$. 于是, 对任意向量 $\boldsymbol{\xi}$, 有

$$\boldsymbol{\xi}=\boldsymbol{\xi}I=\boldsymbol{\eta}+\boldsymbol{\zeta},\quad \boldsymbol{\eta}=\boldsymbol{\xi}k(T)g(T),\quad \boldsymbol{\zeta}=\boldsymbol{\xi}h(T)f(T).$$

因为 $\boldsymbol{\eta}f(T)=\boldsymbol{\xi}k(T)g(T)f(T)=\boldsymbol{\xi}k(T)m(T)=\mathbf{0}$, 类似地, $\boldsymbol{\zeta}g(T)=\mathbf{0}$, 所以这就是所要求的分解.

分解式 (23) 是唯一的, 因为如果 $\boldsymbol{\xi}=\boldsymbol{\eta}_1+\boldsymbol{\zeta}_1=\boldsymbol{\eta}_2+\boldsymbol{\zeta}_2$ 是两个分解式, 那么 $\boldsymbol{\alpha}=\boldsymbol{\eta}_1-\boldsymbol{\eta}_2=\boldsymbol{\zeta}_2-\boldsymbol{\zeta}_1$ 是满足 $\boldsymbol{\alpha}f(T)=\mathbf{0}$ 和 $\boldsymbol{\alpha}g(T)=\mathbf{0}$ 的向量. 因此由 (24), 有

$$\boldsymbol{\alpha}I=\boldsymbol{\alpha}h(T)f(T)+\boldsymbol{\alpha}k(T)g(T)=\mathbf{0}.$$

于是 $\boldsymbol{\eta}_1=\boldsymbol{\eta}_2$, $\boldsymbol{\zeta}_1=\boldsymbol{\zeta}_2$.

定理 17 也可按另一种方式叙述. 子空间 S_1 是由所有满足 $\eta f(T)=\mathbf{0}$ 的向量 $\boldsymbol{\eta}$ 组成, S_2 是由所有满足 $\zeta g(T)=\mathbf{0}$ 的向量 $\boldsymbol{\zeta}$ 组成. 这就是说, S_1 是 $f(T)$ 的零空间, S_2 是 $g(T)$ 的零空间, 此外, 按照 7.8 节中的定义, V 是子空间 S_1 与 S_2 的直和. 这两个子空间, 每一个都被 T 映射到自身, 这样, 按下述一般定义, 它是一个 "不变" 子空间.

向量空间 V 的子空间 S 在线性变换 $T: V\to V$ 之下, 如果由 $\boldsymbol{\xi}\in S$ 推出 $\boldsymbol{\xi}T\in S$, 则称 S 为 T 之下的*不变子空间*. 这时, 对应 $\boldsymbol{\xi}\longmapsto\boldsymbol{\xi}T$ 称为 T 在 S 上的*导出变换*.

显然, 如果 S 在 T 之下是不变子空间, 而 $h(x)$ 是任意多项式, 那么 S 在 $h(T)$ 之下也是不变子空间.

在定理 17 中, 对每个 $\boldsymbol{\eta}\in S_1$, 有 $\boldsymbol{\eta}f(T)=\mathbf{0}$; 因此, 如果 T_1 是 T 在 S_1 上的导出线性变换, 那么 T_1 的极小多项式是 $f(x)$ 的因子 $f_1(x)$. 类似地, S_2 上的导出变换 T_2 的极小多项式是 $g(x)$ 的因子 $g_2(x)$. 所以对任意表示成 (23) 的向量 $\boldsymbol{\xi}$, 有

$$\boldsymbol{\xi}f_1(T)g_2(T)=[\boldsymbol{\eta}f_1(T)]g_2(T)+[\boldsymbol{\zeta}g_2(T)]f_1(T)=\mathbf{0}+\mathbf{0}=\mathbf{0}. \tag{25}$$

因此乘积 $f_1(x)g_2(x)$ 可被极小多项式 $m(x)=f(x)g(x)$ 整除. 因为 f 和 g 是互素的, 这就证明了, $f(x)$ 可整除 $f_1(x)$, $g(x)$ 可整除 $g_2(x)$. 但是 $f_1(x)$ 也可整除 $f(x)$, 所以 $f_1=f$, 同样有 $g_2=g$. 于是我们就得到下面的结果.

定理 17′ *如果 S_1 和 S_2 分别是定理 17 中 $f(T)$ 和 $g(T)$ 的零空间, 那么 V 是 S_1 与 S_2 的直和, 并且 T 在 S_1 与 S_2 上的导出变换 T_1 和 T_2 分别有极小多项式 $f(x)$ 和 $g(x)$.*

可以通过很多方式产生不变子空间. 比如, 设 $f(x)$ 是任意多项式, 那么变换 $f(T): V\to V$ 的值域 —— 即所有向量 $\boldsymbol{\xi}f(T)$(其中 $\boldsymbol{\xi}\in V$) 组成的集合 —— 是在 T 之下的不变子空间, 这是因为 $\boldsymbol{\xi}f(T)T=(\boldsymbol{\xi}T)f(T)$ 也是在这个值域中. 一类特殊的不变子空间是由一个向量生成的循环子空间. 我们现在给出定义.

已知变换 $T: V\to V$ 和 V 中一个向量 $\boldsymbol{\alpha}$, 显然, V 的任意子空间, 如果它包含 $\boldsymbol{\alpha}$, 并且在 T 之下是不变子空间, 那么它一定包含 $\boldsymbol{\alpha}$ 的由 T 的多项式 $f(T)$ 作用而得的所有变换式 $\boldsymbol{\alpha}f(T)$. 但是, 所有这种变换式的集合 $Z_{\boldsymbol{\alpha}}$ 是一个包含 $\boldsymbol{\alpha}$ 的不变子空间, 我们称它是由 $\boldsymbol{\alpha}$ 生成的 *T- 循环子空间*.

现在考虑 $\boldsymbol{\alpha}$ 在 T 的逐次幂之下的一列变换式 $\boldsymbol{\alpha}=\boldsymbol{\alpha}I, \boldsymbol{\alpha}T, \boldsymbol{\alpha}T^2,\cdots$. 显然, 存在第一个与它前面的变换式线性相关的变换式 $\boldsymbol{\alpha}T^d$. 那么有

$$\boldsymbol{\alpha}T^d+c_{d-1}\boldsymbol{\alpha}T^{d-1}+\cdots+c_0\boldsymbol{\alpha}I=\boldsymbol{\alpha}m_{\alpha}(T)=\mathbf{0}, \tag{26}$$

这里 $\boldsymbol{\alpha},\boldsymbol{\alpha}T,\cdots,\boldsymbol{\alpha}T^{d-1}$ 是线性无关的. 于是 $m_{\boldsymbol{\alpha}}(x)=x^d+c_{d-1}x^{d-1}+\cdots+c_0$ 是变

换 $T_{\boldsymbol{\alpha}}$ 的极小多项式, 其中 $T_{\boldsymbol{\alpha}}$ 是 T- 循环子空间 $Z_{\boldsymbol{\alpha}}$ 上的导出变换; 多项式 $m_{\boldsymbol{\alpha}}(x)$ 称为 $\boldsymbol{\alpha}$ 的 T- 阶. 注意, T 把 $Z_{\boldsymbol{\alpha}}$ 的基向量 $\boldsymbol{\alpha}, \boldsymbol{\alpha}T, \cdots, \boldsymbol{\alpha}T^{d-1}$ 中的每一个向量都映射到它的下一个向量, 而 $\boldsymbol{\alpha}T^{d-1}$ 除外, 它被映射到

$$\boldsymbol{\alpha}T^d = -c_0\boldsymbol{\alpha} - c_1\boldsymbol{\alpha}T - \cdots - c_{d-1}\boldsymbol{\alpha}T^{d-1}. \tag{27}$$

对于 $Z_{\boldsymbol{\alpha}}$ 的这组基 $\boldsymbol{\alpha}, \boldsymbol{\alpha}T, \cdots, \boldsymbol{\alpha}T^{d-1}$, 表示 $T_{\boldsymbol{\alpha}}$ 的矩阵的行是基向量的变换式的坐标 $(0, 1, 0, \cdots, 0)$, $(0, 0, 1, \cdots, 0)$, $\cdots$, $(0, \cdots, 0, 1)$,$(-c_0, \cdots, -c_{d-1})$. 这个矩阵恰恰是多项式 $m_{\boldsymbol{\alpha}}(x)$ 的友矩阵, 所以我们就证明了

定理 18 *T 在 T- 循环子空间 $T_{\boldsymbol{\alpha}}$ 上导出的具有 T- 阶 $m_{\boldsymbol{\alpha}}(x)$ 的导出变换可以用 $m_{\boldsymbol{\alpha}}(x)$ 的友矩阵来表示.*

反过来, n 次首一多项式 f 的友矩阵 C_f 表示变换 $T = T_{C_f} : F^n \to F^n$, 它把 F^n 的每个单位向量 $\boldsymbol{\varepsilon}_i$ 变换到下一个单位向量 $\boldsymbol{\varepsilon}_{i+1}$, 把最后一个单位向量 $\boldsymbol{\varepsilon}_n$ 变换到 $\boldsymbol{\varepsilon}_1T^n$. 因此, 像在 (19) 式那样, 整个空间 F^n 是由 $\boldsymbol{\varepsilon}_1$ 生成的 T- 循环子空间, 并具有 T- 阶 $f(x)$.

定理 19 *如果变换 $T: V \to V$ 的极小多项式是 $m(x)$, 那么 V 中每个向量 $\boldsymbol{\alpha}$ 的 T- 阶是 $m(x)$ 的一个因子.*

证明 因为 $m(T) = \boldsymbol{O}, \boldsymbol{\alpha}m(T) = \boldsymbol{0}$, 所以, 由 (26) 式, $m(x)$ 是 $\boldsymbol{\alpha}$ 的 T- 阶 $m_{\boldsymbol{\alpha}}(x)$ 的倍式.

推论 *V 中两个向量 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 张成相同的 T- 循环子空间 $Z_{\boldsymbol{\alpha}} = Z_{\boldsymbol{\beta}}$ 当且仅当 $\boldsymbol{\beta} = \boldsymbol{\alpha}g(T)$, 其中多项式 $g(x)$ 与 $\boldsymbol{\alpha}$ 的 T- 阶 $m_{\boldsymbol{\alpha}}(x)$ 是互素的.*

证明留作习题 (习题 8).

习 题

1. (a) 证明：满足 $\boldsymbol{A}^2 = -\boldsymbol{I}$ 的任意 2×2 实矩阵 $\boldsymbol{A}$ 与矩阵 $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ 相似.

   (b) 证明：不存在 3×3 实矩阵满足 $\boldsymbol{A}^2 = -\boldsymbol{I}$.

   (c) 对于满足 $\boldsymbol{A}^2 = -\boldsymbol{I}$ 的 4×4 实矩阵 $\boldsymbol{A}$, 有什么结果?

2. (a) 证明：满足 $T^2 = T$ 的任意 “幂等” 线性变换 T 的值域和零空间是补子空间 (参看 7.8 节的 (35) 式.)

   (b) 证明：任意两个秩相同的幂等矩阵相似. (*提示*：用 (a) 的结果.)

3. (a) 对所有满足 $\boldsymbol{A}^3 = \boldsymbol{I}$ 的 3×3 复矩阵进行分类.

   (b) 对 3×3 实矩阵解答同样的问题.

4. 每一个平面切变换满足 $\boldsymbol{A}^2 + \boldsymbol{I} = \boldsymbol{A} + \boldsymbol{A}$. 求满足这个方程的 2×2 矩阵的标准型. (*提示*：考虑形式 $\boldsymbol{A} - \boldsymbol{I}$.)

*5. 当标量域具备什么条件时, 具有极小多项式 $x^2 + x - 2$ 的矩阵与 2×2 对角矩阵相似.

6. (a) 在定理 17 中, 证明: $g(T)$ 的值域同 $f(T)$ 的零空间相等.

   (b) 证明: 如果 $f(T)g(T) = \boldsymbol{O}$, 其中 $f(x)$ 和 $g(x)$ 是互素的多项式, 那么, 即使 $f(x)g(x)$ 不是 T 的极小多项式, 定理 17 的结论也成立.
7. 证明: 向量 α 的 T- 阶是满足 $\alpha f(T) = \mathbf{0}$ 的次数最低的首一多项式 $f(x)$.
8. 证明定理 19 的推论.
9. 证明: 已知变换 $T: V \to V$, V 中的向量 α 与 β 具有互素的 T- 阶 $f(x)$ 与 $g(x)$, 那么 $\alpha + \beta$ 具有 T- 阶 $f(x)g(x)$.

*10. 证明: T- 循环空间的每个不变子空间本身是 T- 循环空间. (提示: 考虑循环群相应的性质.)

11. 如果 $f(T) = \boldsymbol{O}$, 而 $f(x)$ 和 $g(x)$ 是互素的, 那么 T 和 $g(T)$ 具有相同的循环子空间.

10.8 第一分解定理

在证明定理 17 和定理 17′ 时用过的构造可以用来分解一般线性变换为 "准素" 分支, 这些分支的极小多项式是不可约多项式的幂. 在这个分解中, k 个子空间的直和的概念起着重要的作用.

定义 我们称向量空间 V 是它的子空间 $S_1, \cdots, S_k$ 的直和 (用符号表示就是 $V = S_1 \oplus \cdots \oplus S_k$), 是指 V 中每个向量 ξ 可以唯一地表示为

$$\xi = \eta_1 + \cdots + \eta_k \quad (\eta_i \in S_i, i = 1, \cdots, k). \tag{28}$$

同 7.8 节的定理 16 完全一样, 我们可以证明

定理 20 如果 V 有子空间 $S_1, \cdots, S_k$, 其中每个 S_i 的维数是 n_i, 并有基 $\alpha_{i1}, \cdots, \alpha_{in_i}$, 那么 V 是 $S_1, \cdots, S_k$ 的直和当且仅当

$$\alpha_{11}, \cdots, \alpha_{1n_1};\ \alpha_{21}, \cdots, \alpha_{2n_2};\ \cdots;\ \alpha_{k1}, \cdots, \alpha_{kn_k} \tag{29}$$

是 V 的一组基.

由此得到, V 的维数是直和被加项 S_i 的维数之和 $n_1 + \cdots + n_k$.

推论 如果 V 是由子空间 $S_1, \cdots, S_k$ 张成, 并且

$$d[V] = d[S_1] + \cdots + d[S_k],$$

那么 V 是 $S_1, \cdots, S_k$ 的直和.

如果空间 V 可以表示成变换 T 之下的真不变子空间的直和, 那么我们称线性变换 $T: V \to V$(或对应于 T 的矩阵) 是完全可约的.

定理 21 如果 V 是不变子空间 $S_1,\cdots,S_k$ 的直和, 已知变换在每个子空间上的导出变换用矩阵 $\boldsymbol{B}_i$ 表示, 那么 T 在 V 上可以用矩阵

$$\boldsymbol{B}=\begin{pmatrix} \boldsymbol{B}_1 & \boldsymbol{O} & \cdots & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B}_2 & \cdots & \boldsymbol{O} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{O} & \boldsymbol{O} & \cdots & \boldsymbol{B}_k \end{pmatrix} \tag{30}$$

表示.

这个矩阵 $\boldsymbol{B}$ 沿对角线排列着矩阵块 $\boldsymbol{B}_1,\cdots,\boldsymbol{B}_k$, 其他位置的元素都是零, 称 $\boldsymbol{B}$ 为矩阵 $\boldsymbol{B}_1,\cdots,\boldsymbol{B}_k$ 的直和. 可以看出, $\boldsymbol{B}$ 的任意多项式 $f(\boldsymbol{B})$ 是 $f(\boldsymbol{B}_1),\cdots,f(\boldsymbol{B}_k)$ 的直和.

证明 对每个不变子空间 S_i, 选取一组基 $\boldsymbol{\alpha}_{i1},\cdots,\boldsymbol{\alpha}_{in_i}$, 所以 $\boldsymbol{B}_i$ 是变换 T 在 S_i 上对于这组基的矩阵表示, 那么把这些基向量合并在一起构成整个空间的一组基 (29). 而且 T 把基向量 $\boldsymbol{\alpha}_{i1},\cdots,\boldsymbol{\alpha}_{in_i}$ 变到第 i 个子空间的向量, 因此变换 T 对于基 (29) 可用上述直和矩阵 (30) 来表示. 证毕

现在考虑在基域上把 T 的极小多项式 $m(x)$ 分解成不同的首一不可约多项式 $p_i(x)$ 的幂的乘积, 写成形式

$$m(x)=p_1(x)^{e_1}\cdots p_k(x)^{e_k},\quad e_i>0. \tag{31}$$

因为不同的 $p_i(x)^{e_i}$ 是互素的, 所以反复运用定理 17′ 就得到

定理 22 如果线性变换 $T:V\to V$ 的极小多项式在基域 F 上可以分解成首一不可约因子 $p_i(x)$ 的乘积 ((31) 式), 那么 V 是不变子空间 $S_i,\cdots,S_k$ 的直和, 其中 S_i 是 $p_i(T)^{e_i}$ 的零空间. T 在 S_i 上的导出变换 T_i 有极小多项式 $p_i(x)^{e_i}$.

这就是我们的第一分解定理. 这些子空间 S_i 称为 V 的在 T 之下的 "准素分支". 它们由 T 唯一确定, 因为分解式 (31) 是唯一的.

一个重要的特殊情形是

推论 元素在 F 中的矩阵 $\boldsymbol{A}$ 在 F 上相似于对角矩当且仅当 $\boldsymbol{A}$ 的极小多项式 $m(x)$ 是 F 上不同线性因子的乘积.

证明 设 $T=T_{\boldsymbol{A}}:F^n\to F^n$ 是对应于 $\boldsymbol{A}$ 的变换. 如果

$$m(x)=(x-\lambda_1)\cdots(x-\lambda_k),\quad \lambda_1,\cdots,\lambda_k\text{是不同的标量}, \tag{32}$$

这个定理表明 V 是空间 $S_1,\cdots,S_k$ 的直和, 其中 S_i 是由满足 $\boldsymbol{\eta}_iT=\lambda_i\boldsymbol{\eta}_i$ 的所有向量 $\boldsymbol{\eta}_i$ 组成, 也就是说, 由属于特征值 λ_i 的所有特征向量组成, S_i 的任意基必由这样的特征向量组成, 所以变换 T 在 S_i 上的矩阵表示是 $\lambda_i\boldsymbol{I}$. 像 (29) 式那样把这些

基合并在一起, 我们可以用一个对角矩阵来表示 T, 这个对角矩阵的对角线上的元素是 $\lambda_1, \cdots, \lambda_k$.

反过来, 如果 $\boldsymbol{D}$ 是任意对角矩阵, 它的不同的对角线元素是 $c_1, \cdots, c_k$, 那么用乘积 $f(\boldsymbol{D}) = (\boldsymbol{D} - c_1\boldsymbol{I}) \cdots (\boldsymbol{D} - c_k\boldsymbol{I})$ 表示的变换把每个基向量映射到0, 因此 $f(\boldsymbol{D}) = \boldsymbol{O}$. $\boldsymbol{D}$ 的极小多项式或者与 $\boldsymbol{D}$ 相似的任意其他矩阵的极小多项式是乘积 $(x - c_1) \cdots (x - c_k)$ 的因子, 因此是不同线性因子的乘积.

习　题

1. 证明定理 20.
2. 在定理 22 中, 设 $q_i(x) = \dfrac{m(x)}{p_i(x)^{e_i}}$, 证明: 那里的子空间 S_i 是 $q_i(T)$ 的值域.

*3. 不用定理 17′, 直接证明定理 22.

4. 证明: 如果 $n \times n$ 矩阵 $\boldsymbol{A}$ 相似于对角矩阵 $\boldsymbol{D}$, 那么 $\boldsymbol{D}$ 的对角线元素 λ_i 出现的次数等于属于特征值 λ_i 的特征向量的集合的维数.
5. 证明: 两个矩阵 $\boldsymbol{B}_1$ 与 $\boldsymbol{B}_2$ 的直和的极小多项式是 $\boldsymbol{B}_1$ 与 $\boldsymbol{B}_2$ 的极小多项式的最小公倍式.
6. 证明: 矩阵 $\boldsymbol{A}$ 的极小多项式能够分解成线性因子当且仅当 $\boldsymbol{A}$ 的特征多项式可以同样地分解.

*7. 设 $\boldsymbol{A}$ 是复矩阵, 它的极小多项式 $m(x) = (x - \lambda_1)^{e_1} \cdots (x - \lambda_r)^{e_r}$ 等于它的特征多项式. 证明: 矩阵 $\boldsymbol{A}$ 同 r 个 $e_i \times e_i$ 三角形矩阵 $\boldsymbol{B}_i$ 的直和相似, $\boldsymbol{B}_i$ 的形状如:

$$\boldsymbol{B}_i = \begin{pmatrix} \lambda_i & 1 & & & 0 \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ 0 & & & & \lambda_i \end{pmatrix}.$$

*8. 证明: 如果 $m(x)$ 是 T 的极小多项式, 那么存在一个向量 $\boldsymbol{\alpha}$, 它的 T- 阶恰好是 $m(x)$. (提示: 利用 10.7 节的习题 9, 首先考虑 $m(x) = p(x)^e$ 的情形, 其中 $p(x)$ 是不可约多项式.)

10.9　第二分解定理

下面我们将指出, 线性变换 $T: V \to V$ 的准素分支 S_i 本身是 T- 循环子空间的直和. 在证明这个命题时, 我们将用到向量空间 V 对于子空间 S 的商空间V/S 的概念. 我们回忆一下 (7.12 节), 商空间 $V' = V/S$ 的元素是 S 的陪集 $\boldsymbol{\xi} + S$, 并且由 $\boldsymbol{\xi}P = \boldsymbol{\xi} + S$ 给出的射影 $P: V \to V/S = V'$ 是一个线性变换. 特别是, 对给定的

$T: V \to V$, 如果子空间 S 是 T 之下的不变子空间, 那么在公式

$$(\boldsymbol{\xi}+S)T' = \boldsymbol{\xi}T + S \tag{33}$$

中, $\boldsymbol{\xi}T+S$ 不依赖于 $\boldsymbol{\xi}'=\boldsymbol{\xi}+S$ 中的代表元 $\boldsymbol{\xi}$ 的选取, 这是因为如果选取另外的表示 $\boldsymbol{\eta}=\boldsymbol{\xi}+\boldsymbol{\zeta}$, 那么由于 $\boldsymbol{\zeta}\in S$, 推出 $\boldsymbol{\zeta}T\in S$, 所以有

$$\boldsymbol{\eta}T+S=\boldsymbol{\xi}T+\boldsymbol{\zeta}T+S=\boldsymbol{\xi}T+S.$$

因此由 (33) 式定义的变换 $T': V' \to V'$ 是单值的; 容易验证 T' 也是线性的. 我们称 T' 是 $V/S=V'$ 上的由 T 导出的变换. 而且对于 T 的任意多项式 $f(T)$, 利用 7.12 节的公式, 由 (33) 得到

$$(\boldsymbol{\xi}+S)f(T') = \boldsymbol{\xi}f(T)+S. \tag{34}$$

特别是, 由 $f(T)=\boldsymbol{O}$ 推出 $f(T')=\boldsymbol{O}'$ 在 V' 中, 所以 V' 中 $\boldsymbol{\xi}'$ 的 T'- 阶整除 V 中 $\boldsymbol{\xi}$ 的 T- 阶.

我们现在准备证明第二分解定理.

定理 23　*如果线性变换 $T: V \to V$ 有极小多项式 $m(x)=p(x)^e$, 它是 V 的标量域 F 上首一不可约多项式 $p(x)$ 的幂, 那么 V 是 T- 循环子空间 Z_i 的直和*

$$V = Z_1 \oplus \cdots \oplus Z_r, \tag{35}$$

其中子空间 $Z_1, Z_2, \cdots, Z_r$ 分别具有 T- 阶

$$p(x)^{e_1},\quad p(x)^{e_2},\cdots,p(x)^{e_r},\quad e=e_1\geqslant e_2\geqslant\cdots\geqslant e_r. \tag{36}$$

V 的任意 T- 循环子空间的直和, 表达式有相同个数的分支子空间和相同的 T- 阶集合 (36).

证明　直和分解的存在性可以通过对 V 的维数 n 用归纳法来证明. 当 $n=1$ 时, V 本身就是循环子空间, 于是直接可得结论.

当 $n>1$ 时, 我们有 $p(T)^e=O$, 而 $p(T)^{e-1}\neq O$, 因此 V 包含一个满足 $\boldsymbol{\alpha}_1 p(T)^{e-1}\neq 0$ 的向量 $\boldsymbol{\alpha}_1$. 所以 $\boldsymbol{\alpha}_1$ 的 T- 阶是 $p(x)^e$, 并且 $\boldsymbol{\alpha}_1$ 生成 T- 循环子空间 Z_1. 因为 Z_1 是 T 之下的不变子空间, 所以 T 导出一个在 $V'=V/Z_1$ 上的线性变换 T'. 因为显然有 $p(T')^e=O'$, 所以在 V' 上 T' 的极小多项式是 $p(x)^e$ 的因子, 我们可以对 $d[V/Z_1]=d[V]-d[Z_1]$ 用归纳法, 把 V/Z_1 分解成 T'- 循环子空间 $Z_2',\cdots,Z_r'$ 的直和, 这些子空间的 T'- 阶是

$$p(x)^{e_2'},\cdots,p(x)^{e_r'},\quad e\geqslant e_2'\geqslant\cdots\geqslant e_r'.$$

引理 1 如果 α_i' 生成 T'- 循环子空间 $Z_i'(i=2,\cdots,r)$, 那么陪集 α_i' 包含代表元 α_i, 而 α_i 的 T- 阶是 α_i' 的 T'- 阶.

证明依赖于下述事实: α_1 的 T- 阶 $p(x)^e$ 是 V 的每个元素的 T- 阶的倍式. 特别设 $p(x)^d$ 是 α_i' 的 T'- 阶, 所以对 α_i' 的任意代表元 $\eta=\eta_i$, $\eta p(T)^d=\alpha_1 f(T)$ 在 α_1 生成的 T- 循环子空间中. 那么

$$\mathbf{0}=\eta p(T)^e=\alpha_1 f(T)p(T)^{e-d}.$$

因为 α_1 有 T- 阶 $p(x)^e$, 所以上式推出 $p(x)^e|f(x)p(x)^{e-d}$, 因此 $f(x)=g(x)p(x)^d$, 其中 $g(x)$ 是某一多项式. 我们现在将证明 $\alpha_i=\eta-\alpha_1 g(T)$ 的 T- 阶 $p(x)^d$ 同 α_i' 的 T'- 阶相等. 这是引理所要求的. 因为 α_i 的 T- 阶是 $\alpha_i'=\alpha_i+Z_1$ 的 T'- 阶 $p(x)^d$ 的倍式, 所以只须注意

$$[\eta-\alpha_1 g(T)]p(T)^d=\eta p(T)^d-\alpha_1 f(T)=\mathbf{0}.$$

证完引理 1, 我们设 Z_i 是由 α_i 生成的 T- 循环子空间. 那么 $d[Z_i]=d[Z_i']$, 这因为这两个维数都等于 α_i' 的公共 T- 阶 $p(x)e_i'$ 的次数. 因此

$$d[V]-d[Z_1]=d[V/Z_1]=d[Z_2]+\cdots+d[Z_r]. \tag{37}$$

通过选取基可以得到, 子空间 $Z_1,\cdots,Z_r$ 张成 V; 因此根据 (37) 和定理 20 的推论得到, V 是直和 $V=Z_1\oplus\cdots\oplus Z_r$, 正如断言所述.

剩下需要证明出现在分解式 (36) 中指数的唯一性, 这只须证明这些指数是由 T 和 V 确定的. 通过对这些子空间维数的计算, 就可做到这一点. 例如, 如果 d 表示 $p(x)$ 的次数, 那么循环子空间 Z_i 的维数是 de_i, 因此整个空间 V 的维数是 $d(e_1+\cdots+e_r)$. 还可以看出, 对任意整数 s, Z_i 在 $p(T)^s$ 之下的像 $Z_ip(T)^s$ 是由 $\beta_i=\alpha_ip(T)^s$ 生成的循环子空间. 当 $e_i>s$ 时, 它的维数是 $d(e_i-s)$, 当 $e\leqslant s$ 时, 它的维数是零.

V 的任意向量 ξ 有唯一的表达式

$$\xi=\eta_1+\cdots+\eta_r\quad(\eta_i\in Z_i,i=1,\cdots,r).$$

因此在 $p(T)^s$ 的值域 $Vp(T)^s$ 中, 任意向量有唯一表达式

$$\xi p(T)^s=\eta_1 p(T)^s+\cdots+\eta_r p(T)^s, \tag{38}$$

其中分量 $\eta_ip(T)^s$ 在空间 $Z_ip(T)^s$ 中. 整数 s 确定一个整数 t, 使得

$$e_1>s,\quad\cdots,\quad e_t>s,\quad e_{t+1}\leqslant s$$

(或者, 当 $e_r > s$ 时, $t = r$). 因此由 (38) 式, $Vp(T)^s$ 是由 $\boldsymbol{\beta}_i = \boldsymbol{\alpha}_i p(T)^s$ 生成的循环子空间 $Z_{\boldsymbol{\beta}_i}(i = 1, \cdots, t)$ 的直和, 并且它的维数是

$$d[Vp(T)^s] = d[(e_1 - s) + \cdots + (e_t - s)]. \tag{39}$$

等式左边的维数由 V 和 T 确定, 它们又依次确定 e_i 如下. 首先取 $s = e - 1 = e_1 - 1$, 则由 (39) 式确定出等于 e 的 e_i 的个数; 其次取 $s = e - 2$, 则由 (39) 式确定出等于 $e - 1$ 的 e_i(如果有的话) 的个数, 等等. 这就证明了指数 $e_1, \cdots, e_r$ 的不变性, 从而完成了定理 23 的证明.

习 题

1. 证明: 如果向量空间 V 是由向量 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 的 T- 循环子空间张成的, 那么 T 的极小多项式是 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 的 T- 阶的最小公倍式.
2. 求 8.5 节习题 3 中的矩阵 $\boldsymbol{B}$ 的极小多项式.
3. 详细证明: 如果变换 $T : V \to V$ 是线性的, 子空间 Z 是 T 之下的不变子空间, 那么 $T' : V/Z \to V/Z$ 是线性的.
4. 根据 (37) 证明: $Z_1, \cdots, Z_r$ 张成空间 V.

10.10 有理标准型与若当标准型

用定理 20 和定理 23 容易得到矩阵在相似变换之下的标准型. 我们只须对于循环子空间上的变换给出标准型.

定理 21 提供了这样的标准型. 如果 $\boldsymbol{A}$ 是任意 $n \times n$ 矩阵, 那么在每个循环子空间中适当选取一组基, 在这个子空间上, $T_{\boldsymbol{A}}$ 用友矩阵表示. 把所有这些基合并起来产生 F^n 的一组基, 对于这组基, $T_{\boldsymbol{A}}$ 可用这些友矩阵的直和来表示. 定理 20 和定理 23 关于唯一性的断言指出, 如此得到的友矩阵集合由 $\boldsymbol{A}$ 唯一确定. 于是我们证明了

定理 24 *元素在域 F 中的任意矩阵 $\boldsymbol{A}$, 在 F 上与一个且只与一个多项式*

$$p_i(x)^{e_{i1}},\ \cdots,\ p_i(x)^{e_{ir_i}},\quad e_{i1} \geqslant \cdots \geqslant e_{ir_i} > 0,\quad i = 1, \cdots, k \tag{40}$$

的友矩阵的直和相似, 这些多项式是首一不可约多项式 $p_1(x), \cdots, p_k(x)$ 的幂, $\boldsymbol{A}$ 的极小多项式是 $m(x) = p_1(x)^{e_{11}} p_2(x)^{e_{21}} \cdots p_k(x)^{e_{k1}}$

这组多项式 (40) 是 $\boldsymbol{A}$ 在相似 (在 F 上) 之下的全系不变式, 称为 $\boldsymbol{A}$ 的*初等因子集合*. 把 $\boldsymbol{A}$ 表示为友矩阵的直和的表达式称为 $\boldsymbol{A}$ 的*准素有理标准型* (用 "准素" 这个词是因为用了不可约多项式的幂, 用 "有理" 这个词是因为分析中只用到域 F 中的有理运算).

推论 1 *$n \times n$ 矩阵 A 的特征多项式是 A 的初等因子之积的 $(-1)^n$ 倍.*

证明 容易看出, 矩阵 $\boldsymbol{B}_1, \cdots, \boldsymbol{B}_q$ 的直和的特征多项式是 $\boldsymbol{B}_1, \cdots, \boldsymbol{B}_q$ 的特征多项式的乘积. 但是根据定理 15, 友矩阵 $\boldsymbol{C}_f$ 的特征多项式, 除了符号外, 就是 $f(x)$. 这两个事实同定理一起证明了推论 1.

推论 2 *方阵的特征值是它的极小多项式的根.*

证明 因为极小多项式 $m(x)$ 可整除特征多项式, 所以极小多项式的任意根都是特征多项式的根, 因此也是特征根 (特征值). 反过来, 根据推论 1, 特征多项式的任意根一定是某个初等因子 $p_i(x)^{e_{ij}}$ 的根, 因此根据定理, 它是 $m(x)$ 的根.

例 具有极小多项式 $(x^2+1)(x+3)^2$ 的任意 6×6 有理矩阵与下列友矩阵直和之一相似:

$$C_{(x^2+1)} \oplus C_{(x^2+1)} \oplus C_{(x+3)^2},$$
$$C_{(x^2+1)} \oplus C_{(x+3)^2} \oplus C_{(x+3)^2},$$
$$C_{(x^2+1)} \oplus C_{(x+3)^2} \oplus C_{(x+3)} \oplus C_{(x+3)}.$$

第一种情形的特征多项式是 $(x^2+1)^2(x+3)^2$; 第二、三种情形的特征多项式都是 $(x^2+1)(x+3)^4$.

在复数域上, 首一不可约多项式只能是线性多项式 $x - \lambda_i$, 其中 λ_i 是标量. 利用这个事实, 对于复矩阵, 或者更一般地, 对于极小多项式是线性因子幂的乘积的任意矩阵, 可以构造出不同的标准型.

这时, 定理 23 中的每个 T- 循环子空间 $Z_{\boldsymbol{\alpha}}$ 将有 T- 阶 $(x-\lambda_i)^e$, 其中 λ_i 为某个标量, e 为正整数. 对于 $Z_{\boldsymbol{\alpha}}$ 的基 $\boldsymbol{\alpha}, \boldsymbol{\alpha}T, \cdots, \boldsymbol{\alpha}T^{e-1}$, 像定理 24 那样, T 可用 $(x-\lambda_i)^e$ 的友矩阵来表示. 另一方面, 考虑向量 $\boldsymbol{\beta}_1 = \boldsymbol{\alpha}, \boldsymbol{\beta}_2 = \boldsymbol{\alpha}U, \cdots, \boldsymbol{\beta}_e = \boldsymbol{\alpha}U^{e-1}$, 其中 $U = T - \lambda_i I$. 因为每个 $\boldsymbol{\beta}_j$ 是 $\boldsymbol{\alpha}T^{j-1}$ 加上向量 $\boldsymbol{\alpha}T^k (k < j-1)$ 的某个线性组合, 所以向量 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_e$ 也是 $Z_{\boldsymbol{\alpha}}$ 的一组基. 为了得到 T 作用在 $\boldsymbol{\beta}_j$ 上的效果, 注意

$$\boldsymbol{\beta}_j T = \boldsymbol{\alpha}U^{j-1}T = \boldsymbol{\alpha}U^{j-1}(U + \lambda_i I) = \lambda_i \boldsymbol{\alpha}U^{j-1} + \boldsymbol{\alpha}U^j.$$

当 $j < e$ 时, 这就得到 $\boldsymbol{\beta}_j T = \lambda_i \boldsymbol{\beta}_j + \boldsymbol{\beta}_{j+1}$; 当 $j = e$ 时, 则有 $\boldsymbol{\alpha}U^j = \boldsymbol{0}, \boldsymbol{\beta}_j T = \lambda_i \boldsymbol{\beta}_j$. 现在对于这组基, T 可用矩阵

$$\begin{pmatrix} \lambda_i & 1 & & & 0 \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ 0 & & & & \lambda_i \end{pmatrix}$$

来表示, 它的行是 $\boldsymbol{\beta}_j T$ 的坐标, 上面列出的矩阵, 主对角线上的元素都是 λ_i, 紧靠主对角线上面的一条对角线上的元素都是 1, 其他元素都是零. 称这样的矩阵为*初等若当 (Jordan) 矩阵*.

如果我们用上述类型的基代替定理 24 中导出友矩阵的那组基, 则我们得到

定理 25 如果矩阵 $\boldsymbol{A}$ 的极小多项式在域 F 上分解成线性因子之积

$$m(x) = (x-\lambda_1)^{e_1}(x-\lambda_2)^{e_2}\cdots(x-\lambda_k)^{e_k}, \tag{41}$$

其中 $\lambda_1,\cdots,\lambda_k$ 是不同的, 那么 $\boldsymbol{A}$ 在域 F 上与一个且只与一个初等若当矩阵的直和相似, 这个直和至少包含一个属于特征根 (特征值)λ_i 的 $e_i\times e_i$ 初等若当矩阵, 并且没有更大的初等若当矩阵属于特征根 (特征值)λ_i.

注意, 在对角线上 λ_i 出现的个数是 λ_i 作为 $\boldsymbol{A}$ 的特征多项式的根的重数.

上面所得到的初等若当矩阵的直和, 如果不计沿对角线排列的这些矩阵块的次序, 它是唯一的, 这种直和称为 $\boldsymbol{A}$ 的*若当标准型*. 它可以应用到复数域上的任意矩阵. 注意, 若当标准型是由初等因子的集合确定的, 特别是, 如果 (41) 中的所有的 e_i 都是 1, 而且只有在这时, 若当标准型成为对角矩阵, 其对角线元素是 $\lambda_i,\cdots,\lambda_k$. 于是, 定理 22 的推论可作为上述定理一种特殊情形.

推论 任意复矩阵同若当标准型矩阵相似.

习 题

1. 在有理数域上, 对下列各矩阵求出所有可能的准素有理标准型:
   (a) 5×5 矩阵, 极小多项式是 $(x-1)^2$.
   (b) 7×7 矩阵, 极小多项式是 $(x^2-2)(x-1)$, 特征多项式是 $(x^2-2)^2(x-1)^3$.
   (c) 8×8 矩阵, 极小多项式是 $(x^2+4)^2(x+8)^2$.
   (d) 6×6 矩阵, 特征多项式是 $(x^4-1)(x^2-1)$.
2. 对具有下列各特征多项式的矩阵, 列出所有可能的若当标准型:
   (a) $(x-\lambda_1)^3(x-\lambda_2)^2$, (b) $(x-\lambda_1)^5(x-\lambda_2)^3$,
   (c) $(x-\lambda_1)(x-\lambda_2)^2(x-\lambda_3)^2$.
3. 把正文中指出的初等若当矩阵表示成准素有理标准型.
4. (a) 证明: 复矩阵和它的转置矩阵一定有相同的若当标准型.
   (b) 推断它们总是相似的.
5. (a) 两个泡利 (Pauli)"旋转矩阵" 满足条件 $\boldsymbol{ST}=-\boldsymbol{TS}, \boldsymbol{S}^2=\boldsymbol{T}^2=\boldsymbol{I}$, 并且是埃尔米特矩阵. 证明: $\boldsymbol{U}=\mathrm{i}\boldsymbol{ST}$ 是埃尔米特矩阵, 并满足 $\boldsymbol{TU}=-\boldsymbol{UT}, \boldsymbol{U}^2=\boldsymbol{I}$.
   (b) 证明: 如果 $\boldsymbol{S}$ 是 2×2 矩阵, 则 $\boldsymbol{S}$ 与 $\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$ 相似, 并且对于这组坐标, $\boldsymbol{T}=\begin{pmatrix}0 & b\\ b^{-1} & 0\end{pmatrix}$, 其中 b 是某数.

*6. 用 10.9 节的方法证明: 任意线性变换 $T:V\to V$ 把 V 分解成具有 T- 阶 $f_1(x),\cdots,f_r(x)$ 的 T- 循环子空间的直和, 这里 $f_i(x)|f_{i-1}(x)(i=2,\cdots,r)$, 而且 $f_1(x)$ 是 T 的极小多项式.

第 11 章　布尔代数与格

11.1　基 本 定 义

我们现在将从近世代数的观点更严密地分析 "集合"(或类) 和 "子集合"的基本概念, 这些概念在 1.11 节中已有过简短介绍. 假设 I 为任意集合, 而 X, Y, Z 表示 I 的子集. 比如 I 是正方形, X, Y, Z 是位于 I 中的全等的互相交叠的圆形, 如图 11-1 的 "维恩 (Venn) 图" 所示.

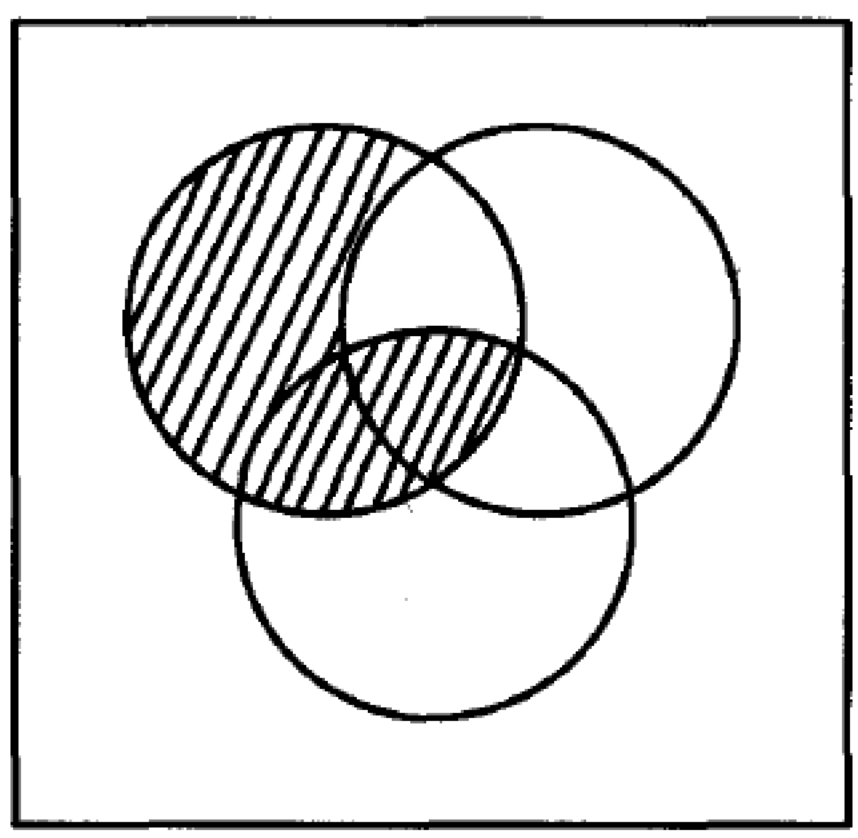


图　11-1

当 X 是 Y 的子集, 即 X 的每个元素都在 Y 中, 我们记作 $X \subset Y$(或 $Y \supset X$). 这个关系也可称为 X"包含" 在 Y 中.

包含关系满足自反律: 这是显然的, 因为任意集合 X 是它本身的子集. 包含关系也满足传递律: 因为如果 X 的每个元素在 Y 中, 并且 Y 的每个元素在 Z 中, 那么显然 X 的每个元素在 Z 中. 但是, 包含关系不满足对称律. 反之, 如果 $X \subset Y$ 且 $Y \subset X$, 那么 X 和 Y 一定包含同样的元素, 因此 $X = Y$.

概括起来, 集合的包含关系与算术的不等关系都具有下列性质:

自反律　对一切 X, 有 $X \subset X$.

反对称律　如果 $X \subset X$ 且 $Y \subset X$, 那么 $X = Y$.

传递律　如果 $X \subset Y$ 且 $Y \subset Z$, 那么 $X \subset Z$.

但是, "对于给定的两个集合 X 和 Y, 不是 $X \subset Y$, 就是 $Y \subset X$" 这个命题是不正确的.

因此两个集合 X 和 Y 的包含关系有四种可能的方式. 一种可能是 $X \subset Y$ 并且 $Y \subset X$, 在这种情况下, 根据反对称律有 $X = Y$. 另一种可能是 $X \subset Y$ 但不满足 $Y \subset X$, 在这种情况下, 我们称 X真包含在Y 中, 并记作 $X < Y$ 或 $Y > X$. 我们还可以有 $Y \subset X$ 但不满足 $X \subset Y$, 在这种情况下, 说 X 真包含 Y. 最后, 我们有既不是 $X \subset Y$, 也不是 $Y \subset X$, 在这种情况下称 X 和 Y 是不可比的. 由于不可比集合的存在, 才使包含关系不同于实数间的不等关系.

已知集合 I 的子集中不仅有包含关系, 而且可以通过两种二元运算 "并" 与 "交" 把它们联系起来, 这两种运算类似于普通的 "加" 与 "乘". 这种类比的程序和重要性首先是由英国数学家布尔 (George Boole, 1815—1864) 发现的, 它在一百多年以前就建立了集合代数的理论.

我们把 X 和 Y 的交(记作 $X \cap Y$) 定义为既在 X 中又在 Y 中的所有元素的集合, 把 X 和 Y 的并(记作 $X \cup Y$) 定义为或者在 X 中或者在 Y 中或者同时在两个集合之中的所有元素的集合. 符号 "$\cap$" 和 "$\cup$" 分别称为 "求交" 运算和 "求并" 运算.

最后, 我们用 X'(读者 "X 的补") 表示不在 X 中的所有的元素的集合. 例如, I' 是空集 $\varnothing$, 它不包含任意元素. 这是因为我们所考虑的只是 I 的子集.

集合的代数运算可以通过图 11-1 的维恩图加以说明. 在这个图中, X, Y, Z 是三个交叠圆形的内部, 这些区域在正方形 I 中的组合可以用适当的阴影区域来表示. 例如, Y' 是 Y 的外部, $X \cap (Y' \cup Z)$ 是图中的阴影区域.

习　题

1. X, Y, Z 的维恩图把正方形分割成八个不交叠的区域, 用 X, Y, Z 的代数组合注明每个这样的区域.
2. 在维恩图上把下列各区域画上阴影:
   $(X' \cap Y) \cup (X \cap Z')$, $(X \cup Y)' \cap Z$, $(Z \cup Y') \cup Z'$.
3. 通过对维恩图上适当的区域画出阴影, 确定下列方程中哪些是成立的:
   (a) $(X' \cup Y)' = X \cap Y'$, (b) $X' \cup Y' = (X \cup Y)'$,
   (c) $(X \cup Y) \cap Z = (X \cup Z) \cap Y$, (d) $X \cup (Y \cap Z)' = (X \cup Y') \cap Z'$.

11.2 定律: 同算术定律类比

我们现在略为详细地描述一下集合代数与普通算术之间的类似, 并用来定义布尔代数, "$\cap, \cup$" 和普通的 "$\cdot, +$" 之间的类似, 由下列定律作部分的描述, 这些定律的正确性是显然的.

幂等律 $X \cap X = X$ 和 $X \cup X = X$.

交换律 $X \cap Y = Y \cap X$ 和 $X \cup Y = Y \cup X$.

结合律 $X\cap(Y\cap Z)=(X\cap Y)\cap Z$ 和 $X\cup(Y\cup Z)=(X\cup Y)\cup Z.$

分配律 $X\cap(Y\cup Z)=(Z\cap Y)\cup(X\cap Z)$ 和 $X\cup(Y\cap Y)=(X\cup Y)\cap(X\cup Z).$

显然, 除了幂等律和第二分配律之外, 所有这些定律都与大家所熟悉的 “·” 与 “+” 的性质相对应, 这些性质在第 1 章已作为公设给出.

下面的基本定律把交和并相互联系起来, 而且把交、并和包含联系起来.

相容律 $X\subset Y, X\cap Y=X$ 和 $X\cup Y=Y$ 这三个条件是互相等价的.

还有, 空集用 $\varnothing$ 表示, $\varnothing$ 和 I 具有下列特殊性质:

泛界 $\varnothing\subset X\subset I$, 对一切 X.

交 $\varnothing\cap X=\varnothing$ 和 $I\cap X=X$.

并 $\varnothing\cup X=X$ 和 $I\cup X=I$.

前三个交和并的性质与普通算术中的 0 和 1 的性质相类似.

最后, 下面三个新的定律把交、并和补联系起来.

互补律 $X\cap X'=\varnothing$ 和 $X\cup X'=I$.

对偶律 $(X\cap Y)'=X'\cup Y'$ 和 $(X\cup Y)'=X'\cap Y'$.

对合律 $(X')'=X$.

如果把 X' 解释为 $1-X$, 并假定 $XX=X$, 那么互补律和对合律与普通算术定律相对应.

上述定律可以用各种方法证明. 第一, 我们可用特殊例子通过 “归纳推理” 来检验它们. 维恩图提供了一个合适的例子. 如果 X 和 Y 分别是图 11-2 中左圆形和右圆形的内部, 那么对于区域 X' 画出水平直线的阴影, 对于区域 Y' 画出垂直直线的阴影. 那么十字阴影线的区域就是 $X'\cap Y'$. 由右图立即看出, 这个区域是并 $X\cup Y$ 的补. 这就是第二对偶律所描述的. 就我们的常识而言, 可以承认这样的论证, 但是, 数学上这是不允许的, 因为在数学推理中, 只允许演绎证明.

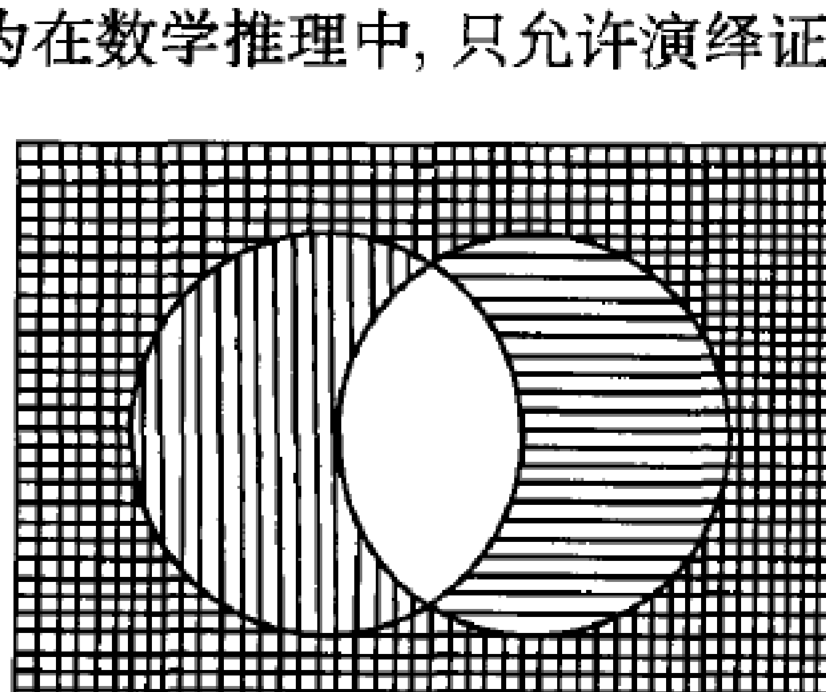


图 11-2

第二, 我们可以把 I 的元素分为四种可能情况来考虑: (i) 元素既在 X 中又在 Y 中; (ii) 元素在 X 中但不在 Y 中; (iii) 元素在 Y 中但不在 X 中; (iv) 元素既不在

X 中也不在 Y 中. 比如, (i) 类元素在 $X\cap Y$ 中因此不在 $(X\cap Y)'$ 中, 不在 $X'\cup Y'$ 中; 而 (ii) 类元素在 $(X\cap Y)'$ 和 Y' 中, 因此在 $X'\cup Y'$ 中, 再看其他两类元素, 也是这个情况. 因此我们看出, $(X\cap Y)'$ 和 $X'\cup Y'$ 具有相同的元素, 这就是第一对偶律. 注意, 对于两个集合 X 和 Y, 元素的四种可能情况用上面这个维恩图的四个区域中的点来表示; 而对于三个集合, 元素有八种情况, 对应于八个区域 (图 11.1).

第三, 我们可以用 "求并" 和 "求交" 运算的叙述性的定义重述这些定律, 例如, 考虑分配律, 这里,

"b 在 $X\cap(Y\cup Z)$ 中" 是指 "b 既在 X 中又在 Y 或 Z 中",

"b 在 $(X\cap Y)\cup(X\cap Z)$ 中" 是指 "b 或者既在 X 中又在 Y 中, 或者既在 X 中又在 Z 中".

按照连词 "既 ……, 又 ……" 及 "或者 ……, 或者 ……" 的通常用法, 稍微 "翻译" 一下就使我们确信这两种叙述是等价的. 分配律的这个证明表明, 集合代数中的定律怎样翻译为 "既 …, 又 …"、"或"、"非" 这些词的性质. 如果我们假定这些性质是基本的, 那么像我们进行平常的数学推理那样, 就能从这些性质证明上述所有关于集合的定律.

习　题

1. 用维恩图验证分配律.
2. 用细分为几种情况的方法验证结合律、交换律和相容律.
3. 利用上述第三种方法中的 "既 ……, 又 ……"、"或"、"非" 等词重述互补律、对偶律和对合律.
4. (a) 在处理四个集合的代数表达式时, 考虑它们元素的所有可能情况, 会出现多少种情况?

   *(b) 画出一个四集合图, 它把元素的每一种可能情况表示成一个区域.

   *(c) 证明: 不存在这样的图形, 在这个图中四个给定的集合都是圆形.
5. 证明: $\varnothing$ 和 I 的交和并的性质可以从泛界性和相容律推导出.
6. 在相容律中, 用 "必要性" 和 "充分性" 代替 "等价" 而得到六个推断.

   证明: 有三个推断对于满足 $0\leqslant x,y\leqslant 1$ 的实数 x,y 是成立的.
7. 在习题 6 中, 如果 $\varnothing$ 用数 0 代替, I 用数 1 代替, 那么 $\varnothing$ 和 I 的交和并的性质中哪些是不成立的? 如果 X' 用 $1-x$ 代替, Y' 用 $1-y$ 代替, 那么关于补的性质中哪些是不成立的?
8. 证明: 对 I 的任意子集 X, Y, $X\subset Y$ 当且仅当 $X'\cup Y=I$.

11.3　布尔代数

我们不再关心由基本逻辑法则来推导上述代数定律, 而是把这些定律中最基本

的定律作为公设 (像第 1 章中的算术定律那样), 然后再从这些公设导出尽可能多的有意义的结论来. 因此, 我们现在用稍微不同的记号给出基本定义, 用这些记号是为了强调这些公设可以用于不同于集合的其他对象.

定义 具有下列性质的元素 $a, b, c, \cdots$ 的集合 B 称为布尔代数.

(i) B 有两个二元运算 $\wedge$(楔形) 和 $\vee$(v 形), 它们满足

幂等律 $a \wedge a = a \vee a = a$,

交换律 $a \wedge b = b \wedge a, a \vee b = b \vee a$,

结合律 $a \wedge (b \wedge c) = (a \wedge b) \wedge c, \qquad a \vee (b \vee c) = (a \vee b) \vee c.$

(ii) 这两个运算满足吸收律

$$a \wedge (a \vee b) = a \vee (a \wedge b) = a.$$

(iii) 这两个运算是互相可分配的

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c), \quad a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c).$$

(iv) B 包含泛界 O, I, 它们满足

$$O \wedge a = O, \quad O \vee a = a, \quad I \wedge a = a, \quad I \vee a = I.$$

(v) B 有求补的一元运算 $a \mapsto a'$, 它遵循下面的互补律

$$a \wedge a' = O, \qquad a \vee a' = I.$$

当然上述所有定律都假定对一切 $a, b, c \in B$ 是成立的.

利用这个定义, 11.1 节和 11.2 节的结论可以概括为下面的命题.

定理 1 在交、并和补三种运算之下, 任意集合 I 的全体子集构成一个布尔代数.

为了更有选择地说明上述这些公设的意义, 我们现在描述几个例子, 在这些例子中, 有一些公设成立, 但不全都成立.

例 1 设 L 是以 n 维欧几里得向量空间 (7.10 节) 的子空间为元素的集合. 这里定义 $S \wedge T = S \cap T$ 是 S 和 T 的交, $S \vee T = S + T$ 是 S 和 T 的线性和, O 是零向量$\mathbf{0}$, I 是整个空间, S' 是子空间 S 的正交补空间 $S^{\perp}$.

那么, 公设 (i), (ii), (iv) 和 (v) 都满足. 但是分配律 (iii) 不满足. (例如, 设 S, T, U 分别是平面上由 (1, 0), (0, 1), (1, 1) 张成的子空间.)

例 2 设 L 是以有限群 G 的正规子群 $M, N, \cdots$ 为元素的集合, 设 $M \wedge N = M \cap N$ 是 M 和 N 的交, 而 $M \vee N = MN$ 是由所有乘积 $xy (x \in M, y \in N)$ 组成的集合. 那么 $M \wedge N$ 和 $M \vee N$ 都是 G 的正规子群. 如果 O 表示群单位元素 1, I 是群 G 本身, 那么虽然 (iii) 和 (v) 一般都不满足, 但公设 (i), (ii) 和 (iv) 都满足.

因为例 1 和例 2 中构造的系统都满足公设 (i) 和 (ii), 所以它们在下述意义下是格.

定义 如果集合 L 有两个二元运算[①] $\wedge$ 和 $\vee$, 它们满足幂等律、交换律和结合律, 并且满足吸收律 (ii), 那么 L 是一个格. 如果除此之外还满足分配律 (iii), 那么 L 称为分配格.

例如, 如果所有多边形区域的集合 L 包含空集 $\varnothing$, 并且面积为零的集合可以忽略不计, 那么集合 L 在交和并运算之下是一个分配格. 再如, 在所有正整数的集合 $\mathbf{Z}^+$ 中, 如果定义 $m \wedge n$ 是 m 和 n 的最大公因子, $m \vee n$ 是 m 和 n 的最小公倍数, 那么 $\mathbf{Z}^+$ 是一个分配格.

上面作为公设的各种定律有很多有趣的代数结论, 我们现在来推导其中最简单的几个.

结合律和交换律的作用已经在 1.5 节中研究过了. 结合律实际上意味着我们可以不用括号来组成多重交或多重并; 交换律意味着, 在只含有 $\vee$ 或只含有 $\wedge$ 的表达式中, 各项可以按照我们喜欢的任何方式排列.

连同上面的定律, 幂等律的作用显然是允许我们消去重复出现的项 —— 只留下一个已知项, 其余重复出现的项全部消去, 概括起来我们有

引理 1 设 f 和 g 是由符号 $\vee$ 和所有字母 $a_1, \cdots, a_n$(可能有些字母重复) 构成的两个表达式, 那么由幂等律、交换律和结合律可推出 $f = g$. 对于只含有 $\wedge$ 的表达式, 上述结论同样成立.

设 N 是下标 $i = 1, 2, \cdots, n$ 的集合, 我们可以不含糊地用

$$\bigvee_N a_i \quad \text{或} \quad \bigvee_{i=1}^n a_i \qquad \text{和} \qquad \bigwedge_N a_i \quad \text{或} \quad \bigwedge_{i=1}^n a_i$$

分别表示所有 a_i 的并(join) 和交(meet). 这些记号类似于代数记号 $\sum$ 和 $\prod$.

再有, 我们从交换律、结合律和分配律出发, 可以用归纳法像 1.5 节那样导出一般分配律如下

$$x\wedge(y_1\vee\cdots\vee y_n) = (x\wedge y_1)\vee\cdots\vee(x\wedge y_n),$$

$$x\vee(y_1\wedge\cdots\wedge y_n) = (x\vee y_1)\wedge\cdots\wedge(x\vee y_n),$$

$$(x_1\vee\cdots\vee x_m)\wedge(y_1\vee\cdots\vee y_n) = (x_1\wedge y_1)\vee(x_1\wedge y_2)\vee\cdots\vee(x_m\wedge y_n).$$

习 题

1. 用归纳法详细证明:

   (a) 在任意分配格中, $x \wedge \bigvee_{i=1}^n y_i = \bigvee_{i=1}^n (x \wedge y_i)$.

   (b) 在任意布尔代数中 $\left(\bigvee_{i=1}^n x_i\right)' = \bigwedge_{i=1}^n x_i'$.

① $\vee$ 运算也称为交 (meet), $\vee$ 运算也称为并 (join), 我们将交替使用这些名称.

*2. 用归纳法详细证明：在任意分配格中,

$$\left(\bigvee_{i=1}^{m} x_i\right) \wedge \left(\bigvee_{j=1}^{n} y_j\right) = \bigvee_{i=1}^{m}\left[\bigvee_{j=1}^{n}\left(x_i \wedge y_j\right)\right].$$

3. 详细证明：当 $n > 1$ 时, 例 1 定义了一个格, 但不是分配格.
4. 证明：当 G 为循环群时, 例 2 定义了一个分配格.
5. 证明：四元素群的所有子群组成的格不是分配格.

11.4 其他基本定律的推导

我们指出, 上面列出的关于布尔代数的公设可推出 11.1 节和 11.2 节中讨论的集合代数的其他基本公式. 例如, 它们可推出 O 和 I 的唯一性, 这些我们并没有假定过.

引理 2 *在任意布尔代数中, 恒等式 $a \wedge x = a$ 和 $a \vee x = x$(对一切 x) 中每一个都可推出 $a = O$. 对偶地有, 恒等式 $a \vee x = a$ 和 $a \wedge x = x$(对一切 x) 中每一个都可推出 $a = I$.*

证明 如果对所有 $x, a \wedge x = a$, 那么特别有 $a \wedge O = a$; 但是由 (iv) 有 $a \wedge O = O$, 因此 $a = O$. 同样, 如果对所有的 $x, a \vee x = x$, 那么 $a \vee O = O$; 但是由 (iv) 有 $a \vee O = a$, 因此又有 $a = O$. I 的唯一性的证明类似.

引理 3 *对任意格中的元素 $a, b, a \wedge b = a$ 成立当且仅当 $a \vee b = b$.*

证明 如果 $a \vee b = b$, 那么根据吸收律 (ii), 有 $a \wedge b = a \wedge (a \vee b) = a$. 反之, 如果 $a \wedge b = a$, 则 $a \vee b = (a \wedge b) \vee b$. 因此根据交换律, $a \vee b = b \vee (b \wedge a) = b$, 这里最后一步又用到 (ii).

推论 *在布尔代数的定义中, 条件 (iv) 可由下列公设中的任何一个来代替：*

(iv′) 对一切 x, 有 $x \wedge O = O$ 和 $x \vee I = I$;

(iv″) 对一切 x, 有 $O \vee x = x$ 和 $I \wedge x = x$.

上面给出的布尔代数的定义没有提到包含关系, 即使包含关系是所有概念中最基本的. 我们现在来定义这个关系, 并由上述公设推导它的基本性质. 其证明重述相容律, 相容律的一部分已经证过了, 如上面引理 3 所述.

定义 *定义 $a \leqslant b$ 是指 $a \wedge b = a$, 或者指 $a \vee b = b$(根据引理 3, 这两个说法是等价的).*

引理 4 *在任意格中, 关系 $a \leqslant b$ 满足自反律、反对称律和传递律.*

证明 因为 $a \wedge a = a$, 所以对一切 a, 有 $a \leqslant a$, 这就证明了自反律, 再有, 由 $a \leqslant b$ 和 $b \leqslant a$ 可推出

$$a = a \wedge b = b \wedge a = b,$$

这就证明了反对称律. 最后, 由 $a \leqslant b$ 和 $b \leqslant c$ 推出 $a = a \wedge b = a \wedge (b \wedge c) = (a \wedge b) \wedge c = a \wedge c$, 因此 $a \leqslant c$. 这就证明了传递律.　　证毕

吸收律的作用在上面引理 2 和引理 3 的证明中已经显示出来. 幂等律在格的定义中是多余的, 实际上, 由吸收律、交换律和结合律可推出幂等律: 因为吸收律就是说, 对所有 x, z, 有 $x = x \wedge (x \vee z)$. 设 $z = x \wedge y$, 我们推出, 对所有 x, y, 有 $x = x \wedge [x \vee (x \wedge y)]$; 再应用对偶的吸收律 $x \vee (x \wedge y) = x$, 于是我们就得到 $x = x \wedge x$(这就是幂等律). $x = x \vee x$ 的证明类似, 只须把 $\wedge$ 和 $\vee$ 互换.

引理 5　*在任意分配格中, 由 $a \vee x = a \vee y$ 和 $a \wedge x = a \wedge y$ 一起可推出 $x = y$.*

证明　通过等式替换, 并逐次应用吸收律和分配律, 我们有

$$\begin{aligned} x &= x \wedge (x \vee a) = x \wedge (y \vee a) \\ &= (x \wedge y) \vee (x \wedge a) = (y \wedge x) \vee (y \wedge a) \\ &= y \wedge (x \vee a) = y \wedge (y \vee a) = y. \end{aligned}$$

现在我们回想一下求补运算 $a \mapsto a'$ 满足

$$a \wedge a' = O \quad 和 \quad a \vee a' = I.$$

但是任意满足 $a \wedge x = O$ 和 $a \vee x = I$ 的元素 x, 根据引理 5, 它一定满足 $x = a'$. 换句话说, 补 a' 由布尔代数定义中的互补律 (v)唯一确定. 我们现在证明补集的其余性质 (对偶律和对合律) 在任意布尔代数中也都成立.

引理 6　*在任意布尔代数中, 我们有*

$$(x')' = x, \qquad (x \wedge y)' = x' \vee y', \qquad (x \vee y)' = x' \wedge y'. \tag{1}$$

证明　"x' 是 x 的补" 这个说法由交换律可推出 "x 是 x' 的补", 这因为 $x' \wedge x = x \wedge x' = O$ 和 $x' \vee x = x \vee x' = I$. 但是我们刚刚证明过补是唯一的, 因此 x 是 x' 的唯一的补, 于是 $(x')' = x$. 再有, 根据分配律有

$$\begin{aligned} (x \wedge y) \wedge (x' \vee y') &= (x \wedge y \wedge x') \vee (x \wedge y \wedge y') \\ &= [(x \wedge x') \wedge y] \vee (x \wedge O) \\ &= [O \wedge y] \vee O = O \vee O = O. \\ (x \wedge y) \vee (x' \vee y') &= (x \vee x' \vee y') \wedge (y \vee x' \vee y') \\ &= (I \vee y') \wedge (y \vee y' \vee x') \\ &= I \wedge (I \vee x') = I. \end{aligned}$$

这就证明了 $x' \vee y'$ 是 $x \wedge y$ 的补. 因此, 再根据补的唯一性, $x' \vee y' = (x \wedge y)'$ 是 $x \wedge y$ 的补. 恒等式 $(x \vee y)' = x' \wedge y'$ 可以类似地证明.

推论　*为了求出由带撇和不带撇的字母通过多重 $\vee$ 和 $\wedge$(但不用加撇的括号) 构成的表达式的补, 可以把表达式中的 $\vee$ 和 $\wedge$ 全都互换, 并把每个不带撇的字母加上撇, 把每个带撇字母的撇去掉.*

例如, 根据这个法则, $(x'\wedge y)\vee(z\wedge w')$ 的补是 $(x\vee y')\wedge(z'\vee w)$.

证明 如果在已知表达式 f 中字母的个数 n(重复的也计算在内) 是 1, 那么推论是正确的, 这因为 $(x)'=x'$, $(x')'=x$. 如果不然, 因为表达式中的括号都不带撇, 所以我们可以把它写成 $f=a\wedge b$ 或 $f=a\vee b$, 由此分别得到 $f'=a'\vee b'$ 或 $f'=a'\wedge b'$. 但是表达式 a 和 b 包含的字母比 f 包含的字母少, 因此, 对 n 用归纳法, 我们可以假定推论对于 a 和 b 都是正确的. 再代入表达式 $f'=a'\vee b'$ 或 $f'=a'\wedge b'$ 中, 我们就得到所要求的补的公式.

习　题

1. 证明：幂等律 $x\vee x=x$ 可由交换律、结合律和吸收律推出.

习题 2 至习题 10 是在布尔代数的条件之下.

2. 详细证明：$(x\vee y)'=x'\wedge y'$.
3. 化简下列布尔表达式：

   (a) $(x'\wedge y')'$,　(b) $(a\vee b)\vee(c\vee a)\vee(b\vee c)$,　(c) $(x\wedge y)\vee(z\wedge x)\vee(x'\vee y')'$.
4. 证明：$(x\wedge y)\vee(x\wedge y')\vee(x'\wedge y)\vee(x'\wedge y')=I$. 利用两个圆的维恩图加以解释.
5. 证明：$x=y$ 当且仅当 $(x\wedge y')\vee(x'\wedge y)=O$.
6. 证明包拉茨基 (Poretzky) 定律：已知 x 和 t, $x=O$ 当且仅当 $t=(x\wedge t')\vee(x'\wedge t)$.
7. (a) 证明：$y\leqslant x'$ 当且仅当 $x\wedge y=O$.　(b) 证明：$y\geqslant x'$ 当且仅当 $x\vee y=I$.
8. 求出下列表达式的补：

   (a) $x\vee y\vee z'$,　(b) $(x\vee y'\vee z')\wedge(x\vee(y\vee z'))$,

   (c) $x\vee(y\wedge(z\vee w'))$,　(d) $(x'\vee y)'\wedge(x\vee y')$,
9. 把引理 6 的推论的论证方法应用到表达式 $(x'\wedge y\wedge z')\vee(x\wedge y')$ 上, 并说明每一步的理由.
10. 证明：$(x\vee y)\wedge(x'\vee z)=(x'\wedge y)\vee(x\wedge z)$.
11. 证明：在任意分配格中, 有

$$(x\wedge y)\vee(y\wedge z)\vee(z\wedge x)=(x\vee y)\wedge(y\vee z)\wedge(z\vee x).$$

12. 具有泛界 O,I 的格 L 中的元素 a, 如果对某个 $x\in L$, 有 $a\wedge x=O$ 和 $a\vee x=I$, 那么称 a 为有补元素. 证明：如果 a 和 b 都是分配格的有补元素, 那么 $a\wedge b$ 和 $a\vee b$ 也都是有补元素.

11.5　布尔多项式的标准型

在前一节里, 我们已经研究了由 $\wedge,\vee$ 和 $'$ 运算构成的各种表达式. 这样的表达式称为 “布尔多项式”(或 “布尔函数”), 显然, 它类似于普通多项式.

我们现在来定义布尔代数 B 的子代数 为 B 中这样的非空子集 S：如果它包含

任意两个元素 x 和 y, 那么它也包含 $x\wedge y, x\vee y, x'$(因而也包含 $O=x\wedge x'$ 和 I). 给定 B 的一个任意非空子集 X, 那么所有值 $p(x_1,\cdots,x_n)$(元素 $x_i\in X$) 组成的集合显然是 B 的包含 X 的最小子代数. 同群的情形一样, 称这个子代数是由 X 生成的. 例如, 任意一个元素 x 生成的子代数由 x, x', O, I 四个元素组成.

这是下面使人感到惊奇的事实的一个特殊情形: 这个事实是 n 个变量 $x_1,\cdots,x_n$ 的不同布尔多项式的个数等于 2^{2^n}. 现在我们来证明它, 以多项式

$$f(x,y,z)=[x\vee z\vee(y\vee z)']'\vee(y\wedge x)$$

为例进行论证.

第一, 如果多项式中任何括号的外边出现撇, 那么总可以应用对偶律 (如 11.4 节的引理 6) 把它移到括号里边. 当所有的撇都移到括号的最里边时, 多项式变成只含有带撇字母和不带撇字母以及作用在它们上面的 $\vee$ 和 $\wedge$ 的表达式. 例如上述例子中,

$$f=[x'\wedge z'\wedge(y\vee z)]\vee(y\wedge x).$$

第二, 如果任意 $\wedge$ 在括号外边, 而括号里包含 $\vee$, 那么根据分配律, $\wedge$ 可以移到括号里边, 像 $c\wedge(a\vee b)=(c\wedge a)\vee(c\wedge b)$ 那样. 结果得到一个多项式, 其中所有的交 $\wedge$ 先组合起来, 然后再接并 $\vee$ 组成, 也就是说, 这个表达式是某些项 $T_1,\cdots,T_k$ 的并, 其中每个 $T_i(i=1,2,\cdots,k)$ 是一些带撇和不带撇字母的交, 在上面例子中,

$$f=(x'\wedge z'\wedge y)\vee(x'\wedge z'\wedge z)\vee(y\wedge x).$$

第三, 某些表达式可以缩短或者略去. 如果字母 "c" 在一项中出现两次, 则可略去一个 "c", 这因为 $c\wedge c=c$. 如果带撇的 c 和不带撇的 c 同时出现在由 $\wedge$ 连接的项中, 那么整个项是 O, 这因为对一切 a, 有 $c\wedge a\wedge c'=O$; 因此这一项在由 $\vee$ 连接的项中可以略去, 因为对一切 b, 有 $O\vee b=b$. 例如上面的例子中,

$$f=(x'\wedge z'\wedge y)\vee(y\wedge x).$$

现在, 如果某一项 T_k 不包含字母 c, 我们可以写成

$$T_k=T_k\wedge I=T_k\wedge(c\vee c')=(T_k\wedge c)\vee(T_k\wedge c'),$$

这里是用两项代替 T_k, 每项中 c 恰好出现一次. 例如在我们的例子中,

$$f=(x'\wedge z'\wedge y)\vee(y\wedge x\wedge z)\vee(y\wedge x\wedge z').$$

最后, 每一项中出现的字母可以重新排列, 使得它们按自然顺序出现. 例如

$$f=(x'\wedge y\wedge z')\vee(x\wedge y\wedge z)\vee(x\wedge y\wedge z')$$

这称为 f 的析取标准型; 于是我们就证明了下面的引理.

引理 任意 $x_i, \cdots, x_n$ 的布尔多项式可以或者化为 O, 或者化为某些项 T_k 的并, 其中 T_k 具有形式

$$T_k = q_1 \wedge q_2 \wedge \cdots \wedge q_n \qquad (\text{每个 } q_j = x_j \text{ 或 } x_j'). \tag{2}$$

也就是说, 可以化为析取标准型.

因为每个 q_j 都有两种可能, 所以我们看到, T_k 恰有 2^n 种可能. 例如, 当 $n = 3$ 时, 任意布尔多项式用我们的方法可以化为 O 或者化为项

$$\begin{gathered} x \wedge y \wedge z, \quad x' \wedge y \wedge z, \quad x \wedge y' \wedge z, \quad x \wedge y \wedge z', \\ x \wedge y' \wedge z', \quad x' \wedge y \wedge z', \quad x' \wedge y' \wedge z, \quad x' \wedge y' \wedge z' \end{gathered} \tag{3}$$

的某一个并. 图 11-1 的三个圆把正方形分成八个区域, 这八个多项式就表示这八个区域, 这个事实并非偶然. 这在几何上意味着, 三个圆 X, Y, Z 的任何布尔组合是图中八个区域的某种选择的并.

像 (3) 中列举的那些基本项称为极小布尔多项式. 换名话说, n 个变量 $x_1, \cdots, x_n$ 的极小布尔多项式 $M(x_1, \cdots, x_n)$是 n 个元素的交 $\bigwedge\limits_{i=1}^{n} q_i$, 其中第 i 个元素 q_i 或者是 x_i 或者是 x_i'. 于是我们就证明了

定理 2 任意已知的 $x_1, \cdots, x_n$ 的布尔多项式或者等于 O 或者等于一组极小多项式 (记为 S) 的并.

现在赋给每个 M 一个 n 位二进数 $\eta(M) = y_1 y_2 \cdots y_n$, 其中数字 y_i 是 1 或 0 应根据上面的 $M = \bigwedge\limits_{i=1}^{n} q_i$ 中 q_i 是 x_i 或 x_i' 而定. 那么函数 $\eta : M \mapsto \eta(M)$ 是由 $x_1, \cdots, x_n$ 的极小多项式的集合到所有 2^n 个 n 位二进数的集合 I 的双射. 例如在 (3) 中, 这些极小多项式所对应的 $\eta(M)$ 值是

$$111, \quad 011, \quad 101, \quad 110, \quad 100, \quad 010, \quad 001, \quad 000.$$

另一方面, $\eta(M)$ 可以看作向量 $\boldsymbol{\eta} = (y_1, y_2, \cdots, y_n) \in \mathbf{Z}_2^n$, 并且可以认为每个布尔多项式 $\bigvee\limits_S M_\eta(x_1, \cdots, x_n)$ 对应于这些向量组成的集合.

如果 $S_i \subset I$ 是由那些第 i 位数字 $y_i = 1$ 的二进数组成, 那么 S_i' 将由那些 $y_i = 0$ 的二进数组成. 因此 (参看下面习题 9), 表示已知极小多项式 $M(S_1, \cdots, S_n)$ 的集合是由单个二进数 $\alpha(M) = a_1 a_2 \cdots a_n$ 组成, 它的第 i 个数字 a_i 是 0 还是 1, 视 S_i 在 M 中带撇还是不带撇而定. 显然, 不同的极小多项式 $M = M_\alpha$ 要用不同的二进数来表示, 因此, M_α 的不同集合中极小多项式的并表示 I 的不同子集, 这就证明了下面的结果.

推论 恰有 2^{2^n} 个不同的 n 变量的布尔函数.

我们现在可以用系统的方法来代替布尔多项式任意的运算. 任意给出的布尔代数中的方程 $E_1 = E_2$, 只要把两边都化为析取标准型就可确定这个方程是正确还是错误.

习　　题

1. 把下列各表达式化为标准型：
(a) $(x \vee y) \wedge (z' \wedge y)'$,　　(b) $(x \vee y) \wedge (y \vee z) \wedge (x \vee z)$.
2. 用把等式两边化为 (析取) 标准型的方法检验下列给出的各方程的正确性.
(a) $[x \wedge (y \vee z)']' = (x \wedge y)' \vee (x \wedge z)$,　(b) $x = (x' \vee y')' \vee [z \vee (x \vee y)']$.
3. 证明：每个布尔多项式具有对偶标准型, 它是某些 "素多项式" 的 "交". 详细描述这些素多项式, 并指出它们是极小多项式的补. 这个结果与多项式分解定理 "域上的每个普通多项式可唯一地表示成不可约多项式的乘积" 有什么类似之处?
4. 用习题 3 的标准型检验习题 2(a) 的方程.
5. 证明：$f(x, y)$ 的标准型是

$$f(x,y) = [f(I,I) \wedge x \wedge y] \vee [f(I,O) \wedge x \wedge y'] \\ \vee [f(O,I) \wedge x' \wedge y] \vee [f(O,O) \wedge x' \wedge y'].$$

6. 证明：任意两个不同的极小多项式的交是 O.
7. 根据一般分配律展开 $I = (x_1 \vee x_1') \wedge \cdots \wedge (x_n \vee x_n')$ 来证明 I 是所有极小布尔多项式的集合的并.
8. 由习题 7 和 $x_i = x_i \wedge I$ 来证明：每个 x_i 是所有那些第 i 项为 x_i 的极小多项式的并.
9. (a) 设 $\bigvee_A M_\alpha$ 表示集合 A 中的所有极小多项式的并, 证明：

$$\left(\bigvee_A M_\alpha\right) \vee \left(\bigvee_B M_\beta\right) = \bigvee_{A \cup B} M_\gamma,\quad \left(\bigvee_A M_\alpha\right) \wedge \left(\bigvee_B M_\beta\right) = \bigvee_{A \cap B} M_\gamma.$$

(b) 证明：如果我们定义极小多项式的空集的并 $\bigvee_\phi M_\alpha$ 是 O, 那么上述公式仍然成立.

*10. 利用习题 7 和习题 9 证明：$(\bigvee_A M_\alpha)' = \bigvee_{A'} m_\alpha$.(提示：运用 11.4 节引理 6.)

*11. 只用习题 8～ 习题 10 独立地证明：每个布尔多项式可以写成极小多项式的并.

11.6 半　　序

前面很少用到 "包含" 的自反律、反对称律和传递律, 然而这些定律是所有定律中最基本的, 因此可以把它们应用到许多非布尔代数的系统中去.

例如, 对于一个集合的所有子集组成的系统, 这些定律显然成立, 这些子集是按照任意特殊性质来划分的 (记作 $\subset$ 或 $\leqslant$). 例如, 这些定律对于任意群的全体子群 (或者全体正规子群!)、任意域的全体子域和任意线性空间的全体子空间, 等等, 都成立 —— 即使所有这些系统都不构成布尔代数. 这些定律对于实数之间的 "小于或等于" 关系 $x \leqslant y$ 和正整数之间的整除关系 $x|y$, 等等, 也都成立.

这些例子暗示了 "半序"这个抽象概念. "半序" 是指任意满足自反律、反对称律和传递律的关系.

定义 一个具有二元关系 $\leqslant$ 的集合 P, 如果这个关系满足自反律、反对称律和传递律, 那么称 P 为半序集.

对于这种类型的任意关系 $a \leqslant b$(读作 "b 包含 a"), 我们可以定义 $a < b$ 的意思是: $a \leqslant b$ 但 $a \neq b$; 而当 $a < b$, 并且没有 x 能满足 $a < x < b$, 这时可称 b覆盖a.

下面引理指出, 任意格可看作一个半序集 (它的完整含义将在下一节里说明).

引理 在任意格中, 如果关系 $x \leqslant y$ 的意思是 $x \wedge y = x$(等价于 $x \vee y = y$), 那么这个关系是一个半序.

具有有限个元素的半序集可以用图方便地表示出来. 系统中的每个元素用小圆圈表示, 如果 $a > b$, 则对应 a 的小圆圈画在对应 b 的小圆圈之上. 然后对于 a 覆盖 b 的情形, 我们从 a 到 b 画一条下降的直线. 我们可以从图上重新构造关系 $a \geqslant b$, 因为 $a > b$ 当且仅当在图中从 b 出发沿着某些上升的直线段爬到 a.

例如, 在图 11-3 中, 第一图表示四元素群的所有子群组成的系统; 第二图表示三点集的所有子集组成的布尔代数; 第三图表示数 1, 2, 4, 8 在整除关系之下组成的系统. 其他几个是随便构造的, 它告诉我们怎样只通过画图就能构造出抽象的半序集. 6.7 节中图 3 是正方形群的所有子群的半序集的图.

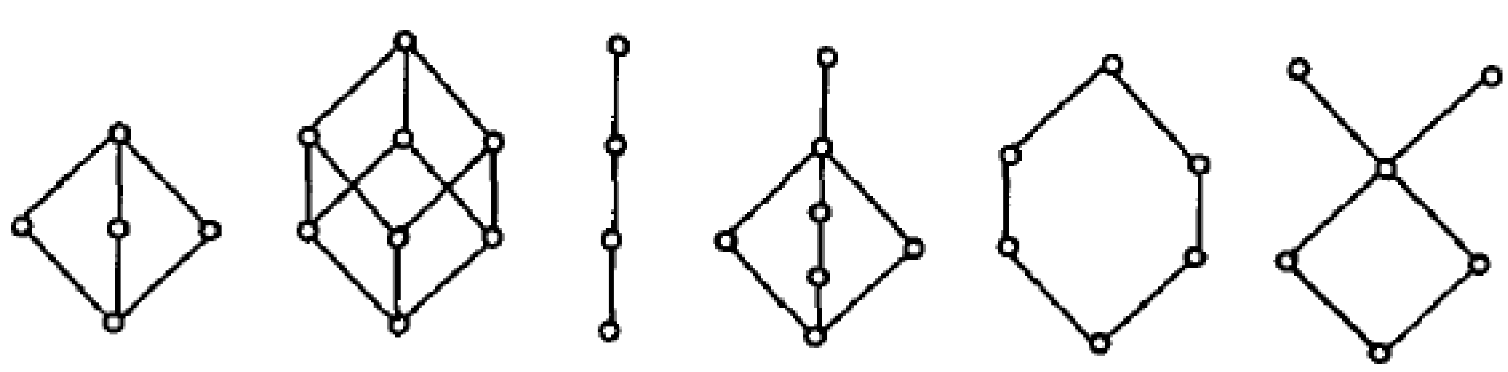


图 11-3

显然, 在任意半序集中, 关系 $\geqslant$ 也满足自反律、反对称律和传递律 (这只不过是从右到左来读这些公设). 因此, 由关系 $a \leqslant b$ 定义的半序集公设能够证明的任何命题, 当每处的 $a \leqslant b$ 用其相反的关系 $a \geqslant b$ 来代替时, 通过一系列同样的推理仍然可以证明它成立. 反之亦然. 这就是

对偶原理 在每个半序集中成立的任何定理, 如果把定理叙述中的所有符号 $\leqslant$ 都与符号 $\geqslant$ 互换, 那么定理仍然成立.

这里要强调指出, 这个原理不是关于半序集的通常意义下的定理, 而是关于定理的定理. 因此, 它属于 "元数学"的范畴.

习 题

1. 详细说明, 图 11-3 中的第二图怎样表示了三点集 I 的所有子集的代数.
2. 画出下列各半序集的图:

   (a) 四点集的所有子集的布尔代数.

(b) 12 阶循环群的所有子群的集合.

(c) 四元数群的所有子集的集合.

(d) 整数 1, 2, 3, 4, 6, 8, 12, 24 在整除性之下构成的半序集.

(e) 54 阶循环群的所有子集的集合.

(f) 模 40 整数的环 $\mathbf{Z}_{40}$ 的所有理想的集合.

3. 证明：习题 2 的 (d), (e), (f) 各部分的半序集在适当规定的意义之下都 "同构".

4. 下列集合中哪些是半序集？

(a) 实数域 $\mathbf{R}$ 的所有子域, 在包含关系之下.

(b) 所有数对 (a,b), 如果 $(a,b) \leqslant (a',b')$ 的意思是 $a \leqslant a'$ 和 $b \leqslant b'$.

(c) 所有实数对 (a,b), 如果 $(a,b) \leqslant (a',b')$ 的意思是或者 $a < a'$ 或者 $a = a'$, 并且 $b \leqslant b'$.

(d) 所有实数对 (a,b), 如果 $(a,b) \leqslant (a',b')$ 的意思是 $a \leqslant a'$ 和 $b \leqslant b'$.

(e) 已知整环的所有子整环, 在包含关系之下.

(f) $F(x)$ 中的所有多项式, 如果 $f(x) \leqslant g(x)$ 的意思是 $f(x)$ 可整除 $g(x)$.

5. 考虑具有关系 $a < b$ 的无素系统, 关系 $a < b$ 满足传递律和非自反律 ($a < a$ 永远不成立). 证明：如果 $a \leqslant b$ 的意思是或者 $a < b$ 或者 $a = b$, 那么集合是半序集.

6. 证明正文中叙述的引理.

7. (a) 证明：11.3 节的例 1 中, 格 L 是通过子空间之间的集合包含关系来定义的.

(b) 叙述并证明：关于 11.3 节例 2 的一个类似的命题.

(c) 如果用 $m|n$ 来定义全体正整数集合的半序, 那么 $\wedge$ 和 $\vee$ 都意味着什么呢？

11.7 格

相容性原理指出怎样通过并和交来定义包含, 现在我们反过来说明, 可通过包含来定义并和交. 也就是说, $x \vee y$ 是既包含 x 又包含 y 的最小集合, 而 $x \wedge y$ 是既包含在 x 中又包含在 y 中的最大集合. 这一说法是由 C.S. 皮尔斯 (Peirce) 提出来的, 我们把它更确切地叙述如下.

设 X 是半序集 P 的某些元素的集合, 如果一个元素 a, 对所有 $x \in X$ 都满足 $a \leqslant x$, 那么称 a 是 X 的 "下界". 像第 4 章所描述的那样, "最大下界"(g.l.b.) 指的是包含其他所有下界的下界, 即最大下界 c, 它对其他任意下界 a, 满足 $c \geqslant a$. 显然, 最大下界如果存在, 就一定唯一, 这是因为如果 a 和 b 都是同一个集合 X 的最大下界, 那么 $a \geqslant b$, 并且 $b \geqslant a$, 因此 $a = b$.

对偶地, 我们可以定义 "上界"和"最小上界"(l.u.b.), 并可证明如果最小上界存在, 就一定唯一. 这里我们正使用了元数学的对偶原理！因此, 我们可以说 "集合的最大下界"、"集合的最小上界" 而不用说 "集合的一个最大下界" 或 "集合的一个最小上界" . 当然这里假定这些界是存在的.

引理 1 在任意格中, 交 $x\wedge y$ 和并 $x\vee y$ 分别是由 x 和 y 两个元素组成的集合的最大下界和最小上界.

证明 因为 $x\wedge x\wedge y=x\wedge y$ 和 $y\wedge x\wedge y=x\wedge y$, 所以相容性原理指出 $x\wedge y$ 是 x 和 y 的下界. 它也是最大下界, 这是因为由 $z\leqslant x$ 和 $z\leqslant y$, 再次根据相容性原理可推出 $z=x\wedge z=x\wedge(y\wedge z)=(x\wedge y)\wedge z$, 所以 $z\leqslant x\wedge y$. 因此 $x\wedge y$ 是最大下界. 由对偶性就完成了整个引理的证明.

这个引理表明, 任意格是具有 "格性质" 的半序集, 所谓 "格性质" 是指任意两个元素具有最大下界和最小上界. 我们现在将指出, 这个性质完全地表征了格.

定理 3 设 L 是任意半序集, 其中任意两个元素 x,y 具有最大下界 $x\wedge y$ 和最小上界 $x\vee y$, 那么在 $\wedge$ 和 $\vee$ 两种运算之下, L 是一个格, 在这个格中, $a\leqslant b$ 当且仅当 $a\wedge b=a$(或等价于 $a\vee b=b$).

证明 只须证明幂等律、交换律、结合律和吸收律及相容性原理. 而且根据对偶原理只须对最大下界来证明幂等律、交换律和结合律. 由定义的对称性, 交换律显然满足. 因为 $x\wedge(y\wedge z)$ 和 $(x\wedge y)\wedge z$ 都是 x,y,z 三个元素的最大下界, 所以满足结合律. 根据定义, 显然 $x\wedge x=x$, 因而幂等律是显然的. 为了证明相容性原理, 首先假定 $x\leqslant y$, 那么任意使得 $z\leqslant x$ 和 $z\leqslant y$ 的 z, 满足 $z\leqslant x$; 而 $x\leqslant x$ 且 $x\leqslant y$, 所以 x 满足最大下界 $x\wedge y$ 的定义, 因此 $x=x\wedge y$. 反之, 如果 $x=x\wedge y$, 那么, x 是 y 的下界, 所以 $x\leqslant y$, 这就证明了相容性原理. 吸收律可通过类似于 11.4 节的引理 3 的证明而推出.

上面没有提到分配律, 因为分配律不是在一切格中都成立. 例如, 当 x,y,z 是四元素群 (图 3 的第一图) 中选出的三个二阶子群时, 分配律就不成立. 然而, 两个与此有关的不等式成立.

定理 4 在任意格中, 半分配律成立:

$$x\wedge(y\vee z)\geqslant(x\wedge y)\vee(x\wedge z),\quad x\vee(y\wedge z)\leqslant(x\vee y)\wedge(x\vee z).$$

此外, 每一个分配律可推出它的对偶形式.

证明 根据对偶原理, 证明可简略一半. 关于第一个半分配律, 请注意, 右边的两项分别是左边两项的下界; 因此 x 和 $y\vee z$ 的最大下界是 $x\wedge y$ 的上界, 又是 $x\wedge z$ 的上界, 所以是 $x\wedge y$ 和 $x\wedge z$ 的最小上界 $(x\wedge y)\vee(x\wedge y)$ 的上界.

最后, 假设 11.3 节 (iii) 中的第一个分配律成立, 我们展开得到

$$\begin{aligned}(x\vee y)\wedge(x\vee z)&=[(x\vee y)\wedge x]\vee[(x\vee y)\wedge z]\\&=x\vee(x\wedge z)\vee(y\wedge z)=x\vee(y\wedge z),\end{aligned}$$

这就是 11.3 节 (iii) 中的另一个分配律. 根据对偶原理, 就完成了定理的证明.

由上面一些定理推得, 要证明一个集合代数是布尔代数, 我们只须知道, (i) 集合包含关系满足自反律、反对称律和传递律; (ii) 两个集合的并是包含这两个集合的最小集合, 两个集合的交是包含在这两个集合中的最大集合; (iii) 恒等地有 $S\cap(T\cup U)=(S\cap T)\cup(S\cap U)$;(iv) 每个集合 S 有一个 "补"S', 满足 $S\cap S'=O, S\cup S'=I$. 这也就证明了

定理 5　*布尔代数是一个分配格, 这个格包含元素 O 和 I, 使得对一切 a, 有 $O\leqslant a\leqslant I$, 并且在这个格中, 每个 a 有一个补 a', 满足 $a\wedge a'=O, a\vee a'=I$.*

布尔代数也可以用许多其他公设系来描述. 例如下面习题 13 就表示了一个布尔代数.

习　　题

1. 图 11-3 中哪些图表示格?
2. 画出两个新图来表示不是格的半序集.
3. 11.6 节的习题 4 中列出的集合中, 哪些表示格?
4. 证明: 如果在格 L 中 $b\leqslant c$, 那么对一切 $a\in L$, 有 $a\wedge b\leqslant a\wedge c$, 和 $a\vee b\leqslant a\vee c$.
5. 叙述并证明布尔代数的对偶原理.
6. 从正文中给出的引理 1 前半部分的证明, 写出后半部分的详细证明, 以此为例说明对偶原理.
7. 证明: 只有有限个元素的格, 具有元素 O 和 I, 它们满足 $O\leqslant x\leqslant I$, 对一切元素 x.

*8. 证明: 含有 O 和 I 的有限半序集, 如果任意适合 $a_i\leqslant b_j(i,j=1,2)$ 的元素 a_1,a_2,b_1,b_2 中间有一个元素 c, 使得对所有的 i 和 j 有 $a_i\leqslant c\leqslant b_j$, 那么这个半序集是一个格.

9. 链是一个全序集 (即是这样的半序集, 其中任意元素 x 和 y 或者满足 $x\geqslant y$, 或者满足 $y\geqslant x$).
   (a) 证明: 每个链是一个分配格.
   (b) 证明: 一个格是链当且仅当它的所有子集合都是子格.

*10. 一个格为模当且仅当由 $x\geqslant z$ 总可推出 $x\wedge(y\vee z)=(x\wedge y)\vee z$.
   (a) 证明: 每个分配格是模.
   (b) 画出表示五元素的格而不是模的图形.
   (c) 证明: 下列各集合是一个模格.
      (i) 一个向量空间的所有子空间.
      (ii) 一个阿贝耳群的所有子群.
      (iii) 任意群的所有正规子群.
   (d) 证明: 在模格中, 由 $x\leqslant z$ 总可推出 $x\vee(y\wedge z)=(x\vee y)\wedge z$. 因此可推断, 对于模格, 对偶原理成立.

*11. 在任意布尔代数中, 两个元素 x 与 y 的对称差定义为 $x+y=(x\wedge y')\vee(x'\wedge y)$.
   (a) 如果 x 和 y 是集合, 它们的对称差指的是什么? 画图说明.

(b) 说明 $x+y$ 满足结合律、交换律, 并且具有零元素.

(c) 证明: 如果把对称差当作和, 把交当作积, 那么每个布尔代数是一个交换环.

12. (a) 证明: 如果我们在向量空间 $\mathbf{Z}_2^n$ 中用 $\boldsymbol{\xi\eta}=(x_1y_1,\cdots,x_ny_n)$ 来定义乘法, 那么 $\mathbf{Z}_2^n$ 变为一个交换环, 在这个环中, 对一切 $\boldsymbol{\xi}$ 有 $\boldsymbol{\xi}^2=\boldsymbol{\xi}$.

*(b) 证明: 在运算 $\boldsymbol{\xi}\wedge\boldsymbol{\eta}=\boldsymbol{\xi\eta}$, $\boldsymbol{\xi}\vee\boldsymbol{\eta}=\boldsymbol{\xi}+\boldsymbol{\eta}+\boldsymbol{\xi\eta}$, $\boldsymbol{\xi}'=(1,1,\cdots,1)-\boldsymbol{\xi}$ 之下, 这个环是一个布尔代数.

*13. 证明: 如果 L 是具有泛界 O 和 I 的格, 在这个格中每个元素 a 都有补 a', 具有性质

$$x\leqslant a' \quad \text{当且仅当} \quad a\wedge x=O,$$

$$y\geqslant a' \quad \text{当且仅当} \quad a\vee y=I.$$

那么 L 是一个布尔代数. (*提示*: 为了证明第一分配律, 只须证明

$$e\equiv[a\wedge(b\vee c)]\wedge[(a\wedge b)\vee(a\wedge c)]'=O.$$

把 e 写成交, 并考虑每一项.)

11.8 集合表示

11.5 节的主要结论是, 对于布尔代数所假定的公设可推出一组恒等式, 这些恒等式对于交、并、补的集合代数来说都是正确的. 实际上, 已经证明了, 一个特殊集合 $\mathbf{Z}_2^n$ 的适当一族子集合 $S_1,\cdots,S_n$ 具有性质: 对于两个布尔多项式 $p,q,p(S_1,\cdots,S_n)=q(S_1,\cdots,S_n)$ 当且仅当它们具有相同的析取标准型. 对于给定的 n, 所有这些析取标准型组成的布尔代数称为具有 n 个生成元的*自由布尔代数*.

我们现在将证明一个更强的结果, 并顺便指出, 用来定义分配格的公设完全地表征了集合的交和并的性质. 为此目的, 我们需要同态和同构的概念, 它们类似于对于群用过的同态和同构概念.

定义 *一个从格 L 到格 M 的函数 $f:L\to M$, 如果对一切 $x,y\in L$, 有 $f(x\wedge y)=f(x)\wedge f(y)$ 和 $f(x\vee y)=f(x)\vee f(y)$, 那么函数 f 称为同态, 一一映上的同态称为同构.*

例如, 由维恩图 (图 11-1) 的三个圆 X,Y,Z 生成的布尔代数与 $\mathbf{Z}_2^3$ 的所有子集组成的代数同构, 相应的函数就像 11.5 节所定义的那样.

引理 1 *两个布尔代数 (看作格) 之间的同构 $f:A\leftrightarrow B$ 一定把 A 中的泛界 O,I 和补映射到 B 中相应的泛界和补.*

证明 显然, "对一切 $x\in A,O\wedge x=O$" 可推出 "对一切 $f(x)\in B,f(O)\wedge f(x)=f(O\wedge x)=f(O)$"; 因此 $f(O)$ 是 B 的泛下界; $f(I)=I$ 的证明类似. 因此 "在 A 中 $x\wedge x'=O$" 可推出 "在 B 中 $f(x)\wedge f(x')=f(x\wedge x')=f(O)=O$", 对偶地有 $f(x)\vee f(x')=I$, 这就证明了 $f(x')=[f(x)]'$, 从而完成了引理 1 的证明.

定义 集环是集合 I 的这样一族子集合：如果这个族包含任意两个集合 S 和 T，那么它一定包含它们的交 $S\cap T$ 和它们的并 $S\cup T$；集域是这样的集环：它包含 I，包含空集 $\varnothing$，并且如果包含任意集合 S，那么也一定包含 S 的补 S'.

换句话说，I 的子集构成的集域恰好是 I 的所有子集构成的布尔代数 A的一个布尔子代数；I 的子集构成的集环恰好是 A 的子格，这时把 A 看作分配格. 我们将证明，每个有限分配格与集环同构，每个有限布尔代数与某 (有限) 集合的所有子集构成的集域同构. 这些结论有点类似于群的凯莱定理.

在证明定理 1 的这些逆命题时，我们还需要下面的概念.

定义 格 L 的一个元素 $a>O$，如果由 $x\vee y=a$ 可推出 $x=a$ 或 $y=a$，则称 a 是并–不可约的；如果 $a<I$，并且由 $x\wedge y=a$ 可推出 $x=a$ 或 $y=a$，则称 a 是交–不可约的；一个元素 p，如果 $p>O$，并且不存在元素 x 使得 $p>x>O$，则称 p 为原子(atom).

引理 2 在布尔代数中，一个元素是并–不可约的当且仅当它是一个原子.

证明 如果 p 是一个原子，那么由 $p=x\vee y$ 推出 $x=p$ 或 $x=O$；在第二种情形中 $p=O\vee y=y$，因此 p 是并–不可约的. 反过来，如果 a 不是原了也不是 O，那么对某个 x 有 $a>x>O$. 因此

$$a=a\wedge I=a\wedge(x\vee x')=(a\wedge x)\vee(a\wedge x')=x\vee(a\wedge x'),$$

这里 $x<a$. 因为 $a\wedge x'\leqslant a$，并且由 $a\wedge x'=a$ 将推出 $x=a\wedge x=a\wedge x'\wedge x=O$，所以还有 $\alpha\wedge x'<a$，因此表明 a 是并–可约的.

现在，对任意有限格 L 的每个元素 a，设 $S(a)$ 是 L 中所有的并–不可约元素 $p_k\leqslant a$ 的集合，考虑映射 $a\mapsto S(a)$. 我们有

引理 3 在有限格 L 中，每个元素 a 满足 $a=\bigvee\limits_{S(a)}p_k$.

证明 对于 $a=O$，可立即得出上述结论. 因为 $S(O)=\varnothing$(空集)，并且 O 是空集的最小上界. 对于任意其他的 $a\in L$，我们应用数学归纳法第二原理，设 $P(n)$ 是命题：当 L 中元素 $x\leqslant a$ 的个数是 $n=n(a)$ 时引理 3 成立. 显然如果 a 是并–不可约的，则 $P(n)$ 正确. 而如果 a 是并–可约的，也不是 O，那么 $a=x\vee y$，其中 $x<a,y<a$，因此 $n(x)<n(a),n(y)<n(a)$. 对 n 用归纳法，由此得到 x 和 y 是并–不可约元素的并：$x=\bigvee\limits_{X}p_\xi$ 和 $y=\bigvee\limits_{Y}q_\eta$，因此 $a=\bigvee\limits_{X}p_\xi\vee\bigvee\limits_{Y}q_\eta$ 是并–不可约元素的并.

引理 4 在任意有限格 L 中，映射 $a\mapsto S(a)$ 把 L 中的交映射到集合论中的交：$S(a\wedge b)=S(a)\cap S(b)$.

证明 根据 $a\wedge b$ 的定义，$P\leqslant a\wedge b$ 当且仅当 $p\leqslant a$ 和 $p\leqslant b$.

引理 5 在有限分配格 L 中，映射 $a\mapsto S(a)$ 把 L 中的并映射到集合论中的并：$S(a\vee b)=S(a)\cup S(b)$.

证明 一个给定的并-不可约元素 p 包含在 $a\vee b$ 中并且仅当

$$p=p\wedge(a\vee b)=(p\wedge a)\vee(p\wedge b).$$

如果 p 是并–不可约的, 则上式意味着或者 $p\wedge a=p$(即 $p\leqslant a$) 或者 $p\wedge b=p$(即 $p\leqslant b$). 这表明, $S(a\vee b)$ 包含 p 当且仅当或者 $S(a)$ 包含 p 或者 $S(b)$ 包含 p. 而反过来显然在任意格中都正确. 证毕

引理 4 和引理 5 表明, 映射 $a\mapsto S(a)$ 是从 L 到集环 $\mathcal{R}$ 上的一个同态, $\mathcal{R}$ 是由 L 的并–不可约元素的集合 I 的所有子集构成. 而且引理 3 表明这个映射是从 L 到 $\mathcal{R}$ 上的一一映射. 这就证明了

定理 6 任意有限分配格 L 与一个集环同构.

当 L 是有限布尔代数时, 引理 2 告诉我们, 每个 $a\in L$ 是全部原子 $p\leqslant a$ 的并. 还有, 根据引理 4 和引理 5, 对任意 $a\in L$, 有

$$S(a)\cap S(a')=S(a\wedge a')=S(O)=\varnothing,$$

$$S(a)\cup S(a')=S(a\vee a')=S(I)=J.$$

这里 J 是 L 中所有原子 (并–不可约元素) 的集合. 这就是 $[S(a)]'=S(a')$, 所以函数 $a\mapsto S(a)$ 是一个同构.

我们已经证明了映射 $a\mapsto T(a)$ 是从任意布尔代数 L 到 L 的原子的子集的集域 $\mathcal{F}$ 上的一个同构. 我们现在指出 $\mathcal{F}$ 包含 L 的所有集合, 从而证明.

定理 7 任意有限布尔代数 L 与它的原子的所有集合组成的布尔代数同构.

证明的完成 这里只剩下证明下面的事实: 如果 S 和 T 是 L 的原子 $p_\sigma,p_\tau,\cdots$ 的两个不同集合, 那么 $\bigvee_S p_\sigma\neq\bigvee_T p_\tau$. 但这是下面引理的推论.

引理 6 如果原子 $q\leqslant\bigvee_S p_\sigma$, 那么 $q\in S$.

因为假定引理 6 成立, 则 $\bigvee_S p_\sigma$ 只包含 S 中的原子, 而不包含其他元素.

引理的证明 根据一般分配律, 有

$$q=q\wedge\bigvee_S p_\sigma=\bigvee_S(q\wedge p_\sigma)$$

因为 q 是并–不可约的, 所以推出上式右边有某一个 $q\wedge p_\sigma=q$, 因此 $O<q\leqslant p_\sigma$, 因为 p_σ 是原子, 这就推出 $q=p_\sigma$.

习　题

1. 证明: 如果两个有限集合 I 和 J 的元素个数相同, 那么 I 的所有子集组成的代数与 J 的所有子集组成的代数同构.

2. 证明：对每个正整数 n, 存在含有 2^n 个元素的布尔代数.
3. 证明：n 元素集合的所有子集组成的布尔代数恰好有 $n!$ 个自同构.
4. (a) 求出一个从布尔代数 A 到布尔代数 B 上的格同态 $f: A \to B$, 这个同态不保持泛界或补.

   (b) 证明：这样的同态保持补当且仅当它保持泛界.
5. (a) 所有正整数的集合 $\mathbf{Z}^+$ 在半序 "$m \leqslant n$ 当且仅当 $m|n$" 之下是一个格.

   (b) 证明：这个格是分配格.

   (c) 鉴定它的并–不可约元素.
6. 证明：如果有限分配格 L 的全体并–不可约元素是链 C, 那么 L 本身是一个链, L 的元素比 C 的元素多多少?

第12章 超限算术

12.1 数与集合

本章将讨论数与集合之间的关系. 这就是对正整数用*基数*的方法处理, 它与 2.6 节中用皮亚诺公设阐明的序的处理方法对比, 后者是把大家所熟悉的序列 "1, 2, 3, 4, ⋯" 看作是基本的. 这种基数处理方法能够使我们按照集合来定义数, 从而减少在数学中必须假定的不加定义的术语的总数. 但是为了实现这一方案, 还需要很大变动的基本概念来与本书吻合.

所以, 我们将假定读者已熟悉正整数和集合的概念, 并从这里进行讨论. 我们的目的是, 推广这种基数方法以便给出*无限基数*的严格定义, 无限基数概念在现代数学中起着很基本的作用. 用这个定义, 我们指出基数如何相加、相乘、并产生任意基数幂, 在这过程中证明, 这些运算具有正整数相应运算的绝大部分 (虽然不是全部) 性质.

数与集合之间的关系来源于下面定义.

定义 *设 n 是任意正数数. 一个集合 S 称为具有基数 n(用符号表示就是 $o(S)=n$*①*) 当且仅当 S 的全体元素与整数 $1,2,3,\cdots,n$ 之间存在双射.*

这个定义意味着 S 的全体元素可以标明 $s_1, s_2, s_3, \cdots, s_n$, 其中 s_k 是 S 中对应于整数 k 的元素. 换句话说, 我们可以来数S 的元素, 一直数到 n, 每个元素数一次而且仅数一次. 由此得到一个推论: 如果两个集合 S 和 T 具有相同的基数, 那么 S 和 T 之间存在一个双射, 即存在对应关系 $s_1 \longleftrightarrow t_1, \cdots, s_n \longleftrightarrow t_n$. 但是下面事实并不显然: 同一个集合不能有两个不同的基数, 也就是说, 按不同顺序重新计数, 我们不能得到不同的元素总数. 我们现在就来证明这个事实, 先叙述一个稍微一般的结果.

定理 1 *设 m 和 n 是正整数. 集合 $\{1,2,\cdots,m\}$ 和集合 $\{1,2,\cdots,n\}$ 的一个真子集之间存在一个双射当且仅当 $m<n$.*

证明 如果 $m<n$, 那么双射: $1 \longleftrightarrow 1, 2 \longleftrightarrow 2, \cdots, m \longleftrightarrow m$ 就是我们所要的对应. 定理 1 的这一半命题比较显然, 但是分析逆命题时必须更谨慎些.

当 $m=1$ 时, 逆命题是显然的, 这因为 1 是最小正整数; 因此我们可以对 m 用归纳法. 我们现在假设集合 $\{1,\cdots,m\}$ 和整数集合 $\{1,\cdots,n\}$ 的真子集 S 之间有一

① 有时称空集为具有零基数的集合.

个双射 $1 \longleftrightarrow f(1), \cdots, m \longleftrightarrow f(m)$. 定义一个新的双射 $i \longleftrightarrow g(i)(i=1,\cdots,m-1)$ 如下：

$$\begin{aligned} g(i) &= f(i), \quad 当\ f(i) \neq n \\ g(i) &= f(m), \quad 当\ f(i) = n \end{aligned} \tag{1}$$

因为至多对一个 i 使 $f(i)=n$, 所以对应 $i \longleftrightarrow g(i)$ 将是整数 $1,\cdots,m-1$ 和整数 $1,\cdots,n-1$ 中的某些整数之间的一一对应.

根据假定, 所有整数 $f(i)$ 的集合 S 是集合 $\{1,\cdots,n\}$ 的**真子集**, 这意味着 S 不包含所有整数 $1,\cdots,n$. 让我们在 S 之外选取一个最小的正整数 $k \leqslant n$, 所以对 $i=1,\cdots,m, f(i)$ 绝不等于 k. 当 $k<n$ 时, 条件 (1) 表明, 没有一个 $g(i)$ 等于 k; 当 $k=n$ 时, $f(i)=n$ 绝不成立, 所以没有一个 $g(i)$ 等于 $f(m)$. 无论哪种情况, 整数 $g(1),\cdots,g(m-1)$ 不能包含所有整数 $1,\cdots,n-1$, 所以 $i \longleftrightarrow g(i)$ 是整数集合 $\{1,\cdots,m-1\}$ 和整数集合 $\{1,\cdots,n-1\}$ 的**真子集**之间的一一对应. 现在根据数学归纳法假设, 我们能得到 $m-1<n-1$, 因此两边都加上 1, 就有 $m<n$.

推论 1　集合 $\{1,\cdots,m\}$ 和集合 $\{1,\cdots,n\}$ 的一个子集之间存在一个双射当且仅当 $m \leqslant n$.

证明　如果 $m \leqslant n$, 那么双射 $1 \longleftrightarrow 1,\cdots,m \longleftrightarrow m$ 就是所要的对应. 反过来, 如果 $i \longleftrightarrow f(i)$ 是集合 $\{1,\cdots,m\}$ 和整数 $1,\cdots,n$ 中某些整数之间的双射, 那么它是集合 $\{1,\cdots,m\}$ 和 $\{1,\cdots,n,n+1\}$ 的真子集之间的双射. 因此根据定理 1, 有 $m<n+1$, 所以 $m \leqslant n$.

推论 2　如果集合 $\{1,\cdots,m\}$ 和集合 $\{1,\cdots,n\}$ 之间存在一个双射, 那么 $m=n$.

这因为, 根据推论 1, 有 $m \leqslant n$ 和 $n \leqslant m$, 因此 $m=n$. 这就表明, 同一个集合不能有两个不同的正整数作为它的基数.

推论 3　如果 S 是集合 $\{1,\cdots,n\}$ 的一个真子集, 那么在集合 $\{1,\cdots,n\}$ 和集合 S 之间不存在双射.

证明　如果存在这样的双射, 由定理 1 将得出 $n<n$, 这是矛盾的.

上述结果可直接推出下面的结论. 设 S 和 T 是任意两个集合, 它们的基数分别为正整数 m 和 n. 那么 $m \leqslant n$ 当且仅当 S 和 T 的子集之间存在一个双射; $m=n$ 当且仅当 S 和整个 T 之间存在一个双射.

习　题

1. 证明：如果集合 S 的基数为 n, t 是 S 中的特定元素, 那么 S 和 $\{1,\cdots,n\}$ 之间存在一个双射, 使得 t 对应于 n.
2. 证明：如果集合 S 的基数为 n, 那么从 S 中去掉一个元素后, 便留下一个基数为 $n-1$ 的集合 S^*.

3. 用证明定理 1 时用过的方法直接证明推论 1.

4. 像习题 3 那样证明推论 3.

12.2 可 数 集

一个集合称为有限集当且仅当它的元素可以用通常的方法计数. 下面我们把这个概念阐述得更确切些.

定义 *一个非空集合 S 称为有限集当且仅当它的基数是一个正整数. 不是空集也不是有限集的集合称为无限集.*

例如, 所有正整数的集合 $\mathbf{Z}^+$ 是无限集. (利用定理 1, 这是不难证明的.) 我们现在引进如下概念: 无限集也可以看作是有基数的.

定义 *一个集合 S, 如果它与所有正整数的集合是双射的, 那么称 S 为可数集, 或称 S 具有基数 $\mathbf{d}$(用符号表示[①]就是 $o(S)=\mathbf{d}$).*

这个定义的条件等价于: 可以把 S 的全部元素列成普通的无限序列 $s_1, s_2, s_3, \cdots, s_n, \cdots$, 使得 S 的每个元素出现一次且仅出现一次. 如果另一个集合 T 与可数集 S 是双射的, 那么可推出 T 本射也是可数的.

定理 2 (伽里略(Galileo) 悖论) *任意可数集都有一个可以把它映射到它的真子集上的双射.*

证明 可数集 (比如说集合 S) 的所有元素可以根据定义写成序列 $s_1, s_2, s_3, \cdots$, 它以全体不同正整数作为下标. 双射 $s_1 \longleftrightarrow s_2, s_2 \longleftrightarrow s_3, \cdots, s_i \longleftrightarrow s_{i+1}, \cdots$ 是集合 S 和从 S 中删去 s_1 而得到的集合之间的一一对应. 证毕

可以看出 $\mathbf{d}$("可数无穷") 是最小的无限基数. 更确切地说, 这是

定理 3 *任意无限集包含一个可数子集.*

证明 设 S 是无限集, 在 S 中选取任意元素 s_1, 然后从 $S-(s_1)$ 中选取第二个元素 s_2; 再从 $S-\{s_1, s_2\}$ 中选取第三个元素 s_3, 等等. 因为 S 是无限的, 所以 $S-\{s_1, s_2, \cdots, s_n\}$ 绝不可能是空集, 因此我们总可以在这里选取一个元素[②] s_{n+1}, 并且这个过程永远不会停止, 直到我们构造出 S 的不同元素的无限序列.

推论 (戴德金–皮尔斯) *一个集合 S 是无限集当且仅当有一个把 S 映射到它的真子集上的双射.*

证明 如果 S 是基数为 n 的有限集, 那么 S 与 $\{1, \cdots, n\}$ 是双射的, 所以定理 1 的推论 3 断言 S 不能与它的一个真子集双射. 相反地, 设 S 是任意无限集, 它将包含由元素 $u_1, u_2, u_3, \cdots$ 组成的可数子集 U. 构造一个函数, 它把 U 中每个元素 u_i 与

① 常常用希伯来字母 $\aleph_0$(读作 "阿勒夫–零") 代替符号 $\mathbf{d}$.

② 这个构造用了集合论中称为选择公理的一个基本原理: 给定任意集合 S, 存在一个 "选择函数" γ, 它从任意非空集合 $T \subset S$ 中选取一个元素 $\gamma(T) \in T$.

其后继 u_{i+1} 对应起来, 再把 S 中不属于 U 的每个元素与其自身对应起来, 这个函数就是从 S 到 S 的真子集的双射. 证毕

实际上, 很多无限集原来都是可数集 (具有基数 $\mathbf{d}$). 下面定理给出两个例子.

定理 4 *所有整数集合 $\mathbf{Z}$ 是可数集; 所有有理数的集合 $\mathbf{Q}$ 是可数集.*

证明 对应 $n \longleftrightarrow 2n+1(n=0,1,2,\cdots),(-n) \longleftrightarrow 2n(n=1,2,3,\cdots)$ 是所有整数的集合 $\{0,-1,1,-2,2,\cdots\}$ 和所有正整数集合 $\{1,2,3,4,5,\cdots\}$ 之间的一一对应. 这就证明了第一个断言.

下面我们来证明, 所有正有理数的集合 $\mathbf{Q}^+$ 是可数集. 为此我们首先把所有正整数的商排成一个无限正方形, 如图 12-1 所示. 如果按照顺序, 围绕较小的正方形边界来排列, 那么我们可以把所有这样的商排成下面普通的无限序列. 第一项是 $\frac{1}{1}$; $\frac{n}{1}$ 的后继是 $\frac{1}{n+1}$; $\frac{m}{n}$ 的后继, 当 $m<n$ 时, 是 $\frac{m+1}{n}$; 当 $m \geqslant n>1$ 时, 是 $\frac{m}{n-1}$. 从这个序列中删去所有不是既约分数的分数 (或者等价地说, 删去的这些分数等于前面已列举过的另一些整数的商). 所得到的子序列把全体正有理数列成一个普通的序列, 并建立起 $\mathbf{Q}^+$ 和 $\mathbf{Z}^+$ 之间的双射 $\frac{m}{n} \longleftrightarrow k$. 而这又可以很容易地扩张成所有有理数的集合 $\mathbf{Q}$ 和所有整数的集合 $\mathbf{Z}$ 之间的双射 $\frac{m}{n} \longleftrightarrow k, 0 \longleftrightarrow 0, -\frac{m}{n} \longleftrightarrow -k$. 因为 $\mathbf{Z}$ 是可数集, 所以推出 $\mathbf{Q}$ 是可数集.

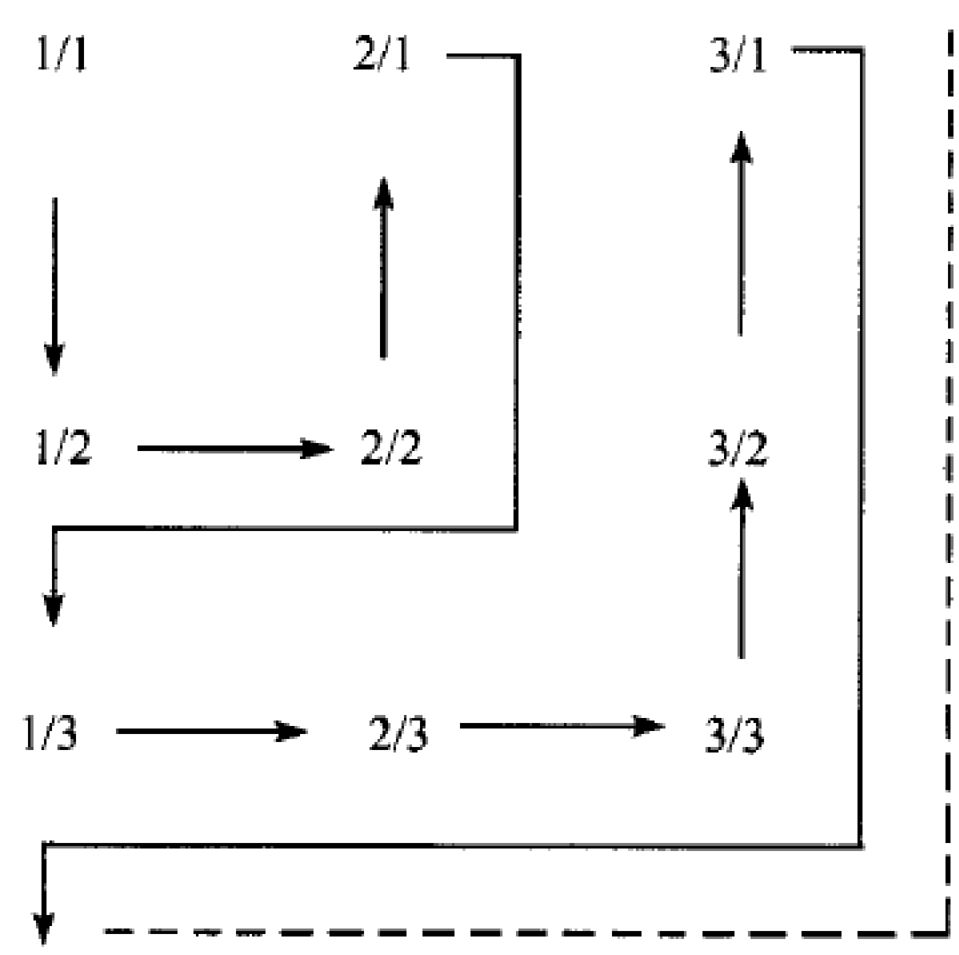


图 12-1

习 题

1. 证明: 7 的所有整数倍数组成的集合是可数集.
2. 证明: 有理数域上的有限维空间 $\mathbf{Q}^n$ 中所有向量组成的集合是可数集.
3. 直接证明: 所有正整数的集合与一个有限集之间的双射是不存在的.
4. 证明: 如果 T 和 U 是可数集, 那么 $S=T \cup U$ 是可数集.

5. 证明：如果 $S=T\cup U$, 其中 S 是可数集, T 是有限集, 那么 U 是可数集.
6. 证明：可数集的每个子集或者是有限集或者是可数集.
7. 证明：仅由数字 9 组成的所有十进小数构成一个无限序列, 这样的十进小数有可数多个.
8. 分别建立所有整数的集合和它的三个真子集之间的特殊的双射.
9. 证明：在图 1 中, 当 $m\leqslant n$ 时, $\frac{m}{n}$ 是第 $(n-1)^2+m$ 项; 当 $m>n$ 时, $\frac{m}{n}$ 是第 m^2-n+1 项.
10. 证明：域 $\mathbf{Q}(\sqrt{2})$ 是可数的 (参看 2.1 节).
11. 证明：每个群包含一个可数子群或有限子群.
12. 列出实数域和它的真子集之间的双射.
13. 证明：所有形为 $r+r'\sqrt{-1}(r,r'\in\mathbf{Q})$ 的数组成的集合是可数的.
*14. 证明：所有有理系数多项式组成的环 $\mathbf{Q}[x]$ 是可数的.

12.3 其他基数

不是所有的无限集合都是可数的：存在不只一个的“无限”基数. 例如

定理 5 (康托 (Cantor)) *所有实数的集合 $\mathbf{R}$ 不是可数的.*

证明 我们应用所谓“对角线法”. 假定所有实数有一个排列 $x_1,x_2,x_3,\cdots$. 把这些数的小数点后的十进小数展开式按它们的排列顺序排成一个正方形阵列, 如图 12-2 所示. 由这个阵列的对角线上的数字按以下方式构造一个新的实数 b(它在 0 和 1 之间)：这里 α_{nn} 是对角线上的第 n 个数字, 设 b_n 是 b 的小数展开式的第 n 个数字, 取

$$x_1=\cdots.a_{11}a_{12}a_{13}a_{14}\cdots$$
$$x_2=\cdots.a_{21}a_{22}a_{23}a_{24}\cdots$$
$$x_3=\cdots.a_{31}a_{32}a_{33}a_{34}\cdots$$
$$x_4=\cdots.a_{41}a_{42}a_{43}a_{44}\cdots$$

图 12-2

$$b_n=\begin{cases}\alpha_{nn}-1, & \text{当 } \alpha_{nn}\neq 0 \text{ 时},\\ 1, & \text{当 } \alpha_{nn}=0 \text{ 时},\end{cases}$$

那么 $b=0.b_1b_2b_3b_4\cdots$ 是某个实数的十进小数 b 的展开式, 这个实数不同于上述排列中的第 n 个数 x_n, 因为至少小数点后第 n 位数字不相等. 于是没有一个 x_n 等于 b, 这与我们的假定“上述排列包含所有实数”相矛盾.

注记 对于下述情形, 这个证明就复杂了：有一些数, 例如 $1.000\cdots=0.999\cdots$, 可以有两种不同的小数展开式, 一种是以无穷个连续的 9 为结尾, 另一种是以无穷个连续的 0 为结尾. 当我们假定原来排列中的小数 $x_1,x_2,\cdots$ 都不用第一类展开式 (带 9 的), 这个麻烦就可以避免. b 的构造永远不会产生数字 $b_n=9$, 因此 b 的小数展开式是与 x_n 的小数展开式进行比较的合适的形式.

定义 *与所有实数的集合 $\mathbf{R}$ 双射的集合 S 称为具有连续统的基数*$\mathbf{c}$(*用符号表示就是* $o(S)=\mathbf{c}$).

实际上，几何学和数学分析中出现的大部分集合具有基数 d 或 c. 可以用几个特殊的结构分别说明这一点，但从长远来看还是先证明施罗德 (E. Schroeder) 和伯恩斯坦 (F. Bernstein) 的一般原理更容易些. 阐述这个原理时，包含着基数的一般概念，我们现在就来定义它.

定义 集合 S 的基数是所有可以双射到 S 的集合 (set) 组成的类[①] (class); S 的基数用 $o(S)$ 表示.

由此得出，两个集合 S 和 T 具有相同基数(或基数等价) 当且仅当 S 和 T 之间有双射. 我们用符号等式 $o(S) = o(T)$ 来表示这个结论.

由于 12.1 节最后一句话，基数之间的不等式概念可以与普通正整数之间不等式的概念一致.

定义 当集合 S 和集合 T 之间有一个由 S 到 T 的单射时，我们就称集合 T 基数优于集合 S，并且记作 $o(S) \leqslant o(T)$.

定理 6 (施罗德 - 伯恩斯坦) 如果 $o(S) \leqslant o(T)$，且 $o(T) \leqslant o(S)$，那么 $o(S) = o(T)$.

换句话说就是，如果存在一个由 S 到 T 的单射，还存在另一个由 T 到 S 的单射，那么就存在整个 S 和整个 T 之间的双射 (逆命题是显然的).

证明 设 $s \longmapsto s\tau$ 是已知的由 S 到 T 的单射，且设 $t \longmapsto t\sigma$ 是已知的由 T 到 S 的子集的单射. S 的每个元素 s 至多是 T 的一个元素 $t = s\sigma^{-1}$ 的像 $t\sigma$; 这个元素 t(如果存在的话) 本身在 S 中也至多有一个像源 $s' = t\tau^{-1} = s\sigma^{-1}\tau^{-1}$，等等. 用这个方法尽可能地追溯 S 的每个元素的 "祖先" (并且对 T 的每个元素也这样做)，我们看出有三种可能情况：(a) 类元素，它的 "祖先" 可以无止境地追溯下去，也许是周期地追溯下去 (见习题 13); (b) 类元素，是 S 中的 "无母祖先" 传下来的; (c) 类元素，是由它在 T 中的 "无母祖先" 传下来的. 对应着这三种情况，我们把 S 分成子集合 S_a, S_b, S_c，把 T 分成子集合 T_a, T_b, T_c. 而且，包含 S 或 T 的任意元素的类一定包含它的 "祖先" 和 "后裔".

实际上，σ 显然是 (τ 也是!)S_a 和 T_a 之间的双射，这因为，S_a 的每个元素是 T_a 的一个元素且只一个元素在 σ 之下的像，而 T_a 的每个元素 t 是 S_a 的一个元素且只一个元素 $t\sigma$ 的像源. 类似地，τ 是 (但 σ 不是!)S_b 和 T_b 之间的双射，而 σ 是 (但 τ 不是!)S_c 和 T_c 之间的双射. 把这三个双射合在一起，$S_a \longleftrightarrow T_a, S_b \longleftrightarrow T_b, S_c \longleftrightarrow T_c$，我们就得到整个 S 和整个 T 之间的双射. 证毕

这些情况可用图 12-3 来解释，但这个图没有标出 "祖先" 可以无止境追溯下去的那些元素. 集合 S 和 T 用两条垂直直线上的点来表示，这里 τ 用朝右斜下方的箭头来表示，σ 用朝左斜下方的箭头来表示. 而 S_b 到 T_b 的双射用不带箭头的直线来表示.

① 这个概念好像 "化学元素" 的概念，同样是抽象概念，化学元素指的是具有一个确定核电荷 (即具有一个确定结构) 的所有原子.

定理 7 直线段 $S_1: 0 < x < 1$, 平面上的单位正方形 $S_2: 0 < x, y < 1$, 空间中的单位立方体 $S_3: 0 < x, y, z < 1$, 都具有基数 **c**.

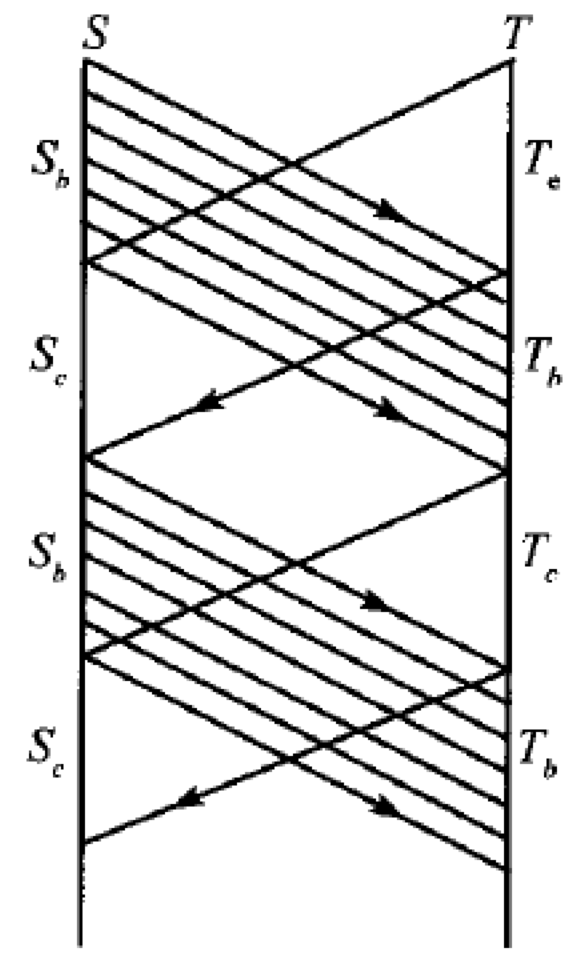


图 12-3

证明 函数 $x \longmapsto \mathrm{e}^x = y$(其逆为 $y \longmapsto \ln y$) 是 $-\infty < x < +\infty$ 和 $0 < y < +\infty$ 之间的一一对应; 函数 $y \longmapsto \dfrac{y}{1+y} = z\left(\text{其逆为 } z \longmapsto \dfrac{z}{1-z}\right)$ 是 $0 < y < +\infty$ 和 $0 < z < 1$ 之间的一一对应. 因此函数 $x \longmapsto \dfrac{\mathrm{e}^x}{1+\mathrm{e}^x} = z$ 是由 $-\infty < x < +\infty$ 到 $0 < z < 1$ 的双射. 这就证明了第一个结论.

为了证明第二个结论, 我们考虑映射

$$(0.x_1x_2x_3\cdots, 0.y_1y_2y_3\cdots) \longmapsto 0.x_1y_1x_2y_2x_3y_3\cdots, \tag{2}$$

这是写成小数形式的 0, 1 间的实数有序对和 0, 1 间单个实数之间的映射. 它是正方形 S_2 和直线段 S_1 的一个子集合之间的单射 (虽然不连续)—— 如果不排除某一位后仅由数字 9 组成的小数, 那么这个映射就是 S_2 和整个 S_1 之间的双射了. 这就证明了 $o(S_2) \leqslant o(S_1)$. 但是又有 $o(S_1) \leqslant o(S_2)\left(\text{通过明显的映射 } x \longmapsto \left(x, \frac{1}{2}\right)\right)$; 因此根据定理 6 得到 $o(S_2) = o(S_1)$, 这就是说 S_2 的基数是 **c**. 类似的映射

$$(0.x_1x_2x_3\cdots, 0.y_1y_2y_3\cdots, 0.z_1z_2z_3\cdots) \longmapsto (0.x_1y_1z_1x_2y_2z_2\cdots) \tag{3}$$

可证明 $o(S_3) \leqslant o(S_1)$, 因此类似地有 $o(S_3) = \mathbf{c}$. 证毕

基数为 **c** 的集合的另一些例子将在习题中给出.

习　题

1. 为什么施罗德–伯恩斯坦定理在集合 T_c 是空集的情况下是显然的? 在这种情况下 S_c 怎么样?
2. 当 S 是区间 $-1 \leqslant s \leqslant \frac{1}{2}$, T 是区间 $-1 \leqslant t \leqslant \frac{1}{2}$, τ 是单射 $s \mapsto s^3$, σ 是双射 $t \mapsto t^3$ 时, 明显地确定集合 $S_a, S_b, S_c, T_a, T_b, T_c$.
3. 当 S 是正整数集合, T 是非负整数集合, τ 是映射 $s \mapsto s$, σ 是映射 $t \mapsto t+1$, 确定集合 $S_a, S_b, S_c, T_a, T_b, T_c$.
4. 证明: n 维实空间中任意包含连续弧的子集合都具有基数 **c**.
5. 证明: 如果存在一个从集合 S 到整个第二个集合 T 的满射, 那么 $o(T) \leqslant o(S)$.
6. 证明习题 5 的逆命题: 如果 $o(T) \leqslant o(S)$, 那么存在一个 S 到 T 上的满射. (你可以假定选择公理成立.)
7. 基数等价关系 $(o(S) = o(T))$ 满足自反律、对称律和传递律吗? 关系 $o(S) \leqslant o(T)$ 满足这些定律吗? 给出证明.

8. 证明：如果 $o(S) \leqslant o(T)$ 且 $o(U) = o(S)$, 那么 $o(U) \leqslant o(T)$.

*9. 证明：以四元数为元素的 $n \times n$ 矩阵共有 $\mathbf{c}$ 个.

*10. 在 0 和 1 之间所有实数的集合与所有无限小数 $0.a_1a_2a_3\cdots$ 的集合之间建立一个明显的双射.

*11. 建立区间 $0 < x < 1$ 和 $0 \leqslant x \leqslant 10$ 之间的双射.

*12. 不用施罗德–伯恩斯坦定理直接证明：

(a) 所有非负实数的集合具有基数 $\mathbf{c}$.

(b) 那些非整数的正实数组成的集合具有基数 $\mathbf{c}$.

*13. 设 $\mathbf{N}$ 是自然数集合, 当 $S = \mathbf{N} \cup \{u, v, w\}, T = \mathbf{N} \cup \{u, v, w\}, \sigma(n) = n + 1$, 在 $\{u, v, w\}$ 上循环 (即 $\sigma(u) = v, \sigma(v) = w, \sigma(w) = u$); $\tau(n) = (n + 2)$, 并且在 $\{u, v, w\}$ 上恒等 (即 $\tau(u) = u, \tau(v) = v, \tau(w = w)$), 明显地确定集合 $S_a, S_b, S_c, T_a, T_b, T_c$.

*12.4　基数的加法与乘法

无限基数就像有限基数那样可以相加和相乘, 除了消去律以外的所有定律仍然成立.

设 m 和 n 是正整数, 我们可以用下面方法构造一个基数为 $m + n$ 的集合：基数为 m 的集合 S'(比如说集合 $\{1, 2, \cdots, m\}$) 再加上一个基数为 n 的与它不相交的集合 S''(比如说集合 $\{m+1, m+2, \cdots, m+n\}$). 那么并 $S' \cup S''$ 就具有基数 $m + n$. 类似地, 所有数对 (i, j) 的集合 (其中 i 跑遍整数 $1, \cdots, m$, j 跑遍整数 $1, \cdots, n$; 例如, $m \times n$ 矩阵的元素的全体下标) 具有基数 mn. 我们并不去证明这些熟悉的事实, 而要指出这些事实暗示着基数加法和乘法运算可以推广到无限基数, 如下所述.

定义　*设 α 和 β 是任意基数. 则分别具有 α 个和 β 个元素的两个不相交子集的并的基数是 $\alpha + \beta$, 所有数对 (x, y) 组成的集合 (这里, x 跑遍具有 α 个元素的集合, y 跑遍具有 β 个元素的集合) 的基数是 $\alpha\beta$.*

基数加法是单值的, 这是因为如果 S 是两个不相交子集 S' 和 S'' 的并, T 是两个不相交子集 T' 和 T'' 的并, 并且 S' 和 T' 之间, S'' 和 T'' 之间都存在双射, 那么我们可以把这两个双射合并成整个 S 和整个 T 之间的一个双射. 类似地, 基数的乘法也是单值的. 事实上, 普通算术的大部分定律, 如同用于有限数的情形一样, 都可用于无限数的情形.①

定理 8　*基数加法和乘法满足交换律和结合律; 乘法对于加法满足分配律; 1 是单位元素.*

证明　基数加法的交换律和结合律是布尔代数定律的推论. 因为不管对什么集合 S 和 T, 函数 $(x, y) \longmapsto (y, x)$ 是所有数对 $(x, y)[x \in S, y \in T]$ 的集合和所有数对

① 可惜, 根据以下定理 (我们不证明), 这个事实就不那么重要了, 这个定理是说, 任意两个无限基数的和或积只是这两个基数中较大的一个. 而超限取幂 (12.5 节) 更为重要.

$(y,x)[y \in T, x \in S]$ 的集合之间的双射, 由此推出乘法交换律. 显然, 所有 3- 数组 $((x,y),z)[x \in S, y \in T, z \in U]$ 的集合到所有 3- 数组 $(x,(y,z))[x \in S, y \in T, z \in U]$ 的集合存在一个双射, 这里 S, T, U 是任意集合, 由此推出乘法结合律. 最后, 如果 T 和 U 是不相交的, 那么所有数对 $(x,w)[x \in S, w \in T$ 或 $U]$ 的集合的基数显然是 $o(S)[o(T)+o(U)]$; 而所有数对 $(x,y)[x \in S, y \in T]$ 的集合与所有数对 $(x,z)[x \in S, z \in U]$ 的集合的并的基数是 $o(S)o(T)+o(S)o(U)$. 这两个集合之间存在一个明显的双射, 因此这就证明了分配律. 对任意基数 α, 显然有 $1 \cdot \alpha = \alpha$.

定理 9 *加法消去律和乘法消去律对无限基数来说是不成立的.*

证明 定理 2 的证明表明 $\mathbf{d} = \mathbf{d} + 1$. 但由此可推出 $\mathbf{d} + 1 = (\mathbf{d} + 1) + 1 = \mathbf{d} + 2$, 虽然 $1 \neq 2$, 这就说明加法消去律不成立. 再有, 正整数的集合 $\mathbf{Z}^+$ 可以分成不相交的偶数集合和奇数集合, 它们都是可数的, 因此 $\mathbf{d} + \mathbf{d} = \mathbf{d}$. 所以, 根据定理 8, 有 $(1+1)\mathbf{d} = 1 \cdot \mathbf{d}$ 或 $2\mathbf{d} = 1\mathbf{d}$, 还是 $2 \neq 1$. 证毕

实际上, 对于所有无限基数, 方程 $\alpha = \alpha + 1$ 和 $\alpha = \alpha + \alpha$ 成立, 但我们不去证明它们.

由此得出一个推论, 有限基数和无限基数的系统不能被嵌入任何一个具有减法和除法的系统中去. 你能证明它吗?

习 题

1. 详细证明 (用布尔代数): 基数加法满足交换律和结合律.
2. 证明: 对任意无限基数 α, 有 $\alpha = \alpha + 1$. (*提示*: 用定理 3.)
3. 证明: $\mathbf{d} + \mathbf{d} + \mathbf{d} = \mathbf{ddd} = \mathbf{d}$, (*提示*: 见图 1.)
4. (a) 证明: 如果 n 是有限基数, 那么 $\mathbf{d} + n = \mathbf{d}$.
   (b) 同样证明: $n\mathbf{d} = \mathbf{d}$.
5. 不用定理 6 证明: $\mathbf{c} + \mathbf{d} = \mathbf{c}$.
6. 不用 12.5 节证明: $\mathbf{c} + \mathbf{c} = \mathbf{cc} = \mathbf{c}$.
7. 证明: $\mathbf{dc} = \mathbf{c}$.
8. 证明 12.4 节中最后一个命题.
9. (a) 证明: 如果 $x \geqslant \mathbf{d}$, 那么 $x + \mathbf{d} = x$.
   (b) 证明: 如果 $x + \mathbf{d} = \mathbf{c}$, 那么 $x = \mathbf{c}$.

*10. 对于*可数群*G, 考虑关于 G 的可能有限子群 S 的阶的拉格朗日定理 (第 6 章) 的证明.
   (a) 证明: 证明中对 S 的阶不加限制.
   (b) 证明: 在可数群 G 中可存在任意给定有限阶的子群.

*12.5 取 幂

如果 S 和 T 是有限集合, 它们的基数分别为 $m = o(S)$ 和 $n = o(T)$, 那么普通幂 $n^m = o(T)^{o(S)}$ 可以描述为由集合 S 到 T 的函数的个数. 任意这样的对应 $x \longmapsto y$

确定一个函数 $y=f(x)$, 它对每个自变量 $x\in S$, 都赋给一个值 $y\in T$. 为了计算所有不同的抽象函数 f 的个数, 我们注意, S 的第一个元素 x 恰好有 $o(T)$ 种可能的像; 对于每一个这样的像, S 的第二个元素的像 y 有 $o(T)$ 种选择, 等等. 所以所有 $o(S)$ 个元素的像的选法总数是 $o(T)$ 自乘 $o(S)$ 次, 即等于 $o(T)^{o(S)}$.

$o(T)^{o(S)}$ 的这种组合特征可以应用于无限基数.

定义　设 α 和 β 是任意非零基数. 则 β^α 是从由 α 个元素组成的集合到由 β 个元素组成的集合的函数的个数.

这定义了一个单叶运算, 即如果 $\alpha=\alpha'$ 且 $\beta=\beta'$, 那么 $\beta^\alpha=\beta'^{\alpha'}$, 其证明是显然的, 我们把它略去.

定理 10　$\mathbf{c}=2^{\mathbf{d}}$.

证明　0 和 1 之间的每个实数 x 有一个二进数展开式 $0.x_1x_2x_3\cdots$, 它可看作一个无穷序列 $x_1,x_2,x_3,\cdots$, 其中 x_i 等于 0 或 1. 不同的实数 x 和 y 具有不同的展开式 (4.3 节), 因此函数 $f(x)=(x_1,x_2,x_3,\cdots)$ 是一一的. 可是这种序列的个数由定义是下面函数的个数, 这种函数是从可数的定义域 (即, 序列中所有 $\mathbf{d}$ 个位置组成的集合) 到两个元素 (即 0 和 1) 组成的取值域上的函数. 我们推出, 0 和 1 之间的实数至多有 $2^{\mathbf{d}}$ 个, 因此根据定理 7 有 $\mathbf{c}\leqslant 2^{\mathbf{d}}$.

另一方面, 每个仅由数字 3 和 7 组成的无限十进小数也可表示不同的实数, 因此 $2^{\mathbf{d}}\leqslant\mathbf{c}$. 现在利用定理 6, 我们就得到 $\mathbf{c}=2^{\mathbf{d}}$.

定理 11　对于任意基数 α,β 和 γ, 下面的取幂法则成立:

(i) $\alpha^\beta\alpha^\gamma=\alpha^{\beta+\gamma}$,　　(ii) $(\alpha\beta)^\gamma=\alpha^\gamma\beta^\gamma$,

(iii) $(\alpha^\beta)^\gamma=\alpha^{\beta\gamma}$,　　(iv) $\alpha^1=\alpha$ 和 $1^\alpha=1$.

证明　(iv) 中的两个恒等式的证明是显然的. 为了证明恒等式 (i)~(iii), 我们假定 S,T 和 U 分别是具有 α,β 和 γ 个元素的集合, 其中 T 和 U 是不相交的.

(i) 的证明　设 V 是这样的集合, 在 V 中, T 和 U 是互补子集, 我们考虑由集合 V 到集合 S 的函数 $h(v)$. 根据定义, 这样函数的个数等于 $\alpha^{\beta+\gamma}$. 另一方面, 每个这样的函数确定一对无关的函数 $(f(t),g(u))$, 并且每对无关函数 $(f(t),g(u))$ 确定一个这样的函数, 这里 $f(t)$ 是由 T 到 S, $g(u)$ 是由 U 到 S. 由定义, 这种函数对的个数是 $\alpha^\beta\alpha^\gamma$.

(ii) 的证明　考虑函数 $h(u)$, 它对每个 $u\in U$ 赋给一对值 $(s,t)=(f(u),g(u))$, s 和 t 分别在 S 和 T 中取任意值. 由定义可知, 这样函数的个数是 $(\alpha\beta)^\gamma$. 但是它也是函数对 $(f(u),g(u))$ 的个数 $\alpha^\gamma\beta^\gamma$, 其中 $f(u)$ 是从 U 到 S, $g(u)$ 是从 U 到 T.

(iii) 的证明　考虑两个变量 $t\in T$ 和 $u\in U$ 的函数 $f(t,u)$, 它在 S 中取值, 根据定义, 这种函数的个数是 $\alpha^{\beta\gamma}$. 但是, 对每个固定的 u, $f(t,u)$ 联系一个对应关系 $f_u(t)$, 它赋给每个 t 一个值 $f_u(t)=f(t,u)\in S$. 反过来, 每个映射 $u\longmapsto f_u$ 定义了两个变量 t 和 u 的函数 $f(t,u)=f_u(t)$. 因为根据定义 f_u 的个数是 α^β, 所以 $f(t,u)$

的个数是 $(\alpha^{\beta})^{\gamma}$. 证毕

由定理 10 和定理 11, 我们可以从相应的关于 $\mathbf{d}$ 的方程推出一些包含 $\mathbf{c}$ 的方程. 例如,

$$\mathbf{c}^2 = (2^{\mathbf{d}})^2 = 2^{2\mathbf{d}} = 2^{\mathbf{d}} = \mathbf{c},$$
$$2\mathbf{c} = 2^1 2^{\mathbf{d}} = 2^{1+\mathbf{d}} = 2^{\mathbf{d}} = \mathbf{c},$$
$$\mathbf{c}^{\mathbf{d}} = (2^{\mathbf{d}})^{\mathbf{d}} = 2^{\mathbf{d}^2} = 2^{\mathbf{d}} = \mathbf{c} \qquad (\text{参看定理 } 4).$$

利用这些结果和下面的习题 1 以及定理 6, 我们容易地得到像 $\mathbf{d}^{\mathbf{d}} = \mathbf{c}$ 和 $n^{\mathbf{d}} = \mathbf{c}$(对任意自然数 $n > 1$) 等这样的法则.

定理 12 对任意基数 α, 有 $\alpha < 2^{\alpha}$.

说明 这个记号是指 $\alpha \leqslant 2^{\alpha}$, 但 $\alpha \neq 2^{\alpha}$.

证明 设 S 是基数为 α 的任意集合. 则 2^{α} 就是函数 $f(x), g(x), \cdots$ 的个数, 这些函数的定义域是 S, 取值为 0 和 1. 定义 $f_x(y) = 0$, 当 $x \neq y$; $f_x(x) = 1$, 于是我们就得到 S 和从 S 到集合 $\{0,1\}$ 的函数的特殊集合之间的双射 $x \longleftrightarrow f_x$. 这就证明了 $\alpha \leqslant 2^{\alpha}$.

反过来, 设在 S 和以 S 为定义域、取值为 0 和 1 的函数之间, 给定任意双射 $x \longleftrightarrow g_x$. 构造一个新函数 $h(x) : h(x) = 0$, 当 $g_x(x) = 1$; $h(x) = 1$, 当 $g_x(x) = 0$. 这就定义了一个以 S 为定义域、取值为 0 和 1 的函数, 而且由构造可知, 对所有函数 $g_x, h(x) \neq g_x(x)$. 我们得出结论: h 与每个 g_x 都不同, 因此在 S 和以 S 为定义域、取值为 0 和 1 的所有函数的集合之间不存在任何双射. 用符号表示就是, $\alpha \neq 2^{\alpha}$.

习　题

1. 证明: 如果 $\alpha \leqslant \beta$, 那么对所有 γ, 有
   (a) $a + \gamma \leqslant \beta + \gamma$, (b) $\alpha\gamma \leqslant \beta\gamma$, (c) $\alpha^{\gamma} \leqslant \beta^{\gamma}$, (d) $\gamma^{\alpha} \leqslant \gamma^{\beta}$.
2. 证明: $\mathbf{c}^{\mathbf{c}} = 2^{\mathbf{c}}$. (提示: 用 12.4 节习题 7.)
3. 证明: 如果集合 S 的基数为 α, 那么 S 的所有可能子集的集合具有基数 2^{α}. (提示: 每个子集合 $T \leqslant S$ 定义一个特征函数 $f_T(x) : f_T(x) = 1$, 当 $x \in T$; $f_T(x) = 0$, 当 $x \notin T$.)
4. 证明: 正方形的所有子集的个数等于所有实变量实函数的个数.

*5. (a) 有多少个实数的有限集合?
   (b) 有多少个实数的可数集合?

*6. 有多少个实数的集合, 它们的基数是 $\mathbf{c}$?

7. 约定对一切 $\alpha > 0$, 有 $0^0 = 1$ 和 $0^{\alpha} = 0$. 证明: 这个约定与定理 11 的定律 (i)~(iv) 是相容的.

第13章 环与理想

13.1 环

在这一章中, 我们将着手研究一般的环以及它们的同态, 还要指出后者是如何同理想有关. 然后, 我们把理想的概念应用到与代数曲线、代数曲面有关的几何上去, 并且应用到代数数的分解理论中去 (在第 14 章中). 我们的基本公设如下所述.

定义 *环 A 是这样的元素系统. 它在加法运算之下是一个阿贝耳群, 在乘法运算之下是封闭的, 这个乘法满足结合律, 并且对于加法满足分配律. 于是, 对于环 A 中所有的 a, b, c 有*

$$a(bc) = (ab)c,\ a(b+c) = ab + ac,\ (a+b)c = ac + bc. \tag{1}$$

我们还要假定每个环 A 有一个单位元素 $1 \neq 0$, 满足 $1a = a1 = a$, 对一切 $a \in A$.

环包括在第 1 ~ 3 章中所研究的所有整环和其他交换环, 例如 $\mathbf{Z}_m$(模 m 整数) 和 $A[x], A[x,y]$(系数在任意给定的交换环 A 中的多项式环). 它还包括*非交换环*, 例如 8.11 节的四元数环. 任意给定域 F 上的所有 $n \times n$ 矩阵组成的集合 $M_n(F)$ 在 $A+B$ 和 AB 运算之下是一个环, 当 $n > 1$ 时, 它也是非交换环.

如果 A 和 B 是任意两个环, 那么所有数对 (a, b)(其中 $a \in A$, $b \in B$) 组成的集合, 在由

$$\begin{aligned}(a_1, b_1) + (a_2, b_2) &= (a_1 + a_2, b_1 + b_2),\\ (a_1, b_1)(a_2, b_2) &= (a_1a_2, b_1b_2)\end{aligned} \tag{2}$$

定义的两种运算之下是一个环. 这样得到的环称为 A 与 B 的*直和*, 记作 $A \oplus B$. 例如, 设 $\mathbf{Q}$ 是有理数域, $\mathbf{Z}$ 是整数环, Q 是四元数环, 则 $\mathbf{Q} \oplus \mathbf{Z} \oplus Q$ 是一个环. 这个奇妙的例子对很多类型的环给出了某种表示!

交换环理论的大部分可以推广到非交换环上去. 例如, 1.12 节中给出的环同构的定义, 不管是否满足条件 $ab = ba$, 都适用; 3.3 节中给出的子环定义也是如此. 此外, 交换环中的很多讨论都可以应用到任意环上. 例如我们可以证明, 一个环 A 的子集合 S 是子环当且仅当 $1 \in S$, 当 b 和 c 在 S 中时可推出 $b - c$ 和 bc 都在 S 中. 另外一些讨论见习题 1.

线性代数[①] 矩阵和四元数是具有可加向量空间结构的一类环的重要例子. 原先这

① 关于线性代数这部分材料主要是作为例子的来源, 并由于我们对它本身感兴趣, 在这里叙述. 它和 13.6 节可以省略, 这并不影响连贯性.

些环是作为比 $\mathbf{C}$ 更广泛的“结合代数系统”来构造的, 今天通常称它们为线性结合代数.

定义 域 F 上的一个线性代数是一个集合 $\mathfrak{A}$, 它是 F 上的一个有限维向量空间, 并且乘法满足结合律和双线性

$$\boldsymbol{\alpha}(\boldsymbol{\beta}\boldsymbol{\gamma})=(\boldsymbol{\alpha}\boldsymbol{\beta})\boldsymbol{\gamma}, \qquad \text{(结合律)} \quad (3)$$

$$\boldsymbol{\alpha}(c\boldsymbol{\beta}+d\boldsymbol{\gamma})=c(\boldsymbol{\alpha}\boldsymbol{\beta})+d(\boldsymbol{\alpha}\boldsymbol{\gamma}), \quad (c\boldsymbol{\alpha}+d\boldsymbol{\beta})\boldsymbol{\gamma}=c(\boldsymbol{\alpha}\boldsymbol{\gamma})+d(\boldsymbol{\beta}\boldsymbol{\gamma}), \qquad \text{(双线性)} \quad (4)$$

对 F 中的所有标量 c 和 d, 以及 $\mathfrak{A}$ 中所有元素 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$, 上述这些定律都成立. $\mathfrak{A}$ 的阶是把 $\mathfrak{A}$ 看作向量空间时它的维数. 当对 $\mathfrak{A}$ 中所有 $\boldsymbol{\alpha}$ 有 $1\boldsymbol{\alpha}=\boldsymbol{\alpha}=\boldsymbol{\alpha}1$ 时, $\mathfrak{A}$ 有一个单位元素1. 如果再增加一个条件: 对 $\mathfrak{A}$ 中每个 $\boldsymbol{\alpha}\neq\mathbf{0}$, 则 $\mathfrak{A}$ 中有元素 $\boldsymbol{\alpha}^{-1}$, 满足 $\alpha^{-1}\alpha=1$, 那么称这个代数是可除代数.

特别是, 每个线性代数是一个环.

著名的弗罗比尼乌斯 (Frobenius) 定理 (1878) 指出, 全体四元数构成实数域上唯一的非交换可除代数.

例 1 在实数域上构造一个由“对偶数”组成的代数, 它包含两个基元素 δ 和 ε, 它们是按照法则 $\delta\varepsilon=\varepsilon\delta=\delta, \delta^2=0, \varepsilon^2=\varepsilon$ 来相乘. 根据这些法则, 可以求出 A 的任意两个元素的乘积, 因为

$$(a\delta+b\varepsilon)(c\delta+d\varepsilon)=ac\delta^2+ad\delta\varepsilon+bc\varepsilon\delta+bd\varepsilon^2=(ad+bc)\delta+bd\varepsilon.$$

可以验证, 它满足像乘法结合律那样的必要的公设. 这个例子很像四元数集合, 它表明如何通过给出基元素的适当的乘法表就可以定义一个代数.

例 2 域 F 上所有 $n\times n$ 矩阵的矩阵代数 $M_n(F)$, 以矩阵 E_{ij} 作为它的基, E_{ij} 的 (i,j) 位置的元素是 1, 其余的元素都是零. 基元素的乘法表是 $\boldsymbol{E}_{ij}\boldsymbol{E}_{jk}=\boldsymbol{E}_{ik}, \boldsymbol{E}_{ij}\boldsymbol{E}_{kl}=\boldsymbol{O}(j\neq k)$.

例 3 设 G 是含有元素 $\alpha_1,\cdots,\alpha_n$, 并有乘法运算 $\alpha_i\alpha_j=\alpha_k$ 的有限群. 如果 F 是任意域, 那么存在 F 上的线性代数 $\mathfrak{A}$, 它以 G 的元素作为基元素, 在 $\mathfrak{A}$ 中, 根据 G 的群表, 并由双线性来定义乘法,

$$(x_1\alpha_1+\cdots+x_n\alpha_n)(y_1\alpha_1+\cdots+y_n\alpha_n)=\sum_{i,j}(x_iy_j)(\alpha_i\alpha_j).$$

这个代数称为G 在 F 上的群代数.

特别是, 以 α 为生成元的二阶循环群的群代数具有基元素 $1=\alpha^2$ 和 α, 并有乘法

$$(x\cdot 1+y\alpha)(u\cdot 1+v\alpha)=(xu+yv)1+(xv+yu)\alpha.$$

对于基 $\beta=\dfrac{1+\alpha}{2}, \gamma=\dfrac{1-\alpha}{2}$, 这个群代数有乘法表 $\beta^2=\beta, \gamma^2=\gamma, \beta\gamma=\gamma\beta=0$.

例 4　$2n \times 2n$ 矩阵中, 在右上方和左下方的 $n \times n$ 矩阵块都是零矩阵的所有矩阵组成的集合构成一个代数, 这个代数是 $M_{2n}(F)$ 的一个子环. 它是两个 $M_n(F)$ 的直和.

我们现在证明关于矩阵的类似于凯莱定理 (6.5 节定理 8) 的一个定理. 同一个域 F 上的两个代数 $\mathfrak{A}$ 和 $\mathfrak{A}'$, 当它们的元素之间存在双射 $\alpha \longleftrightarrow \alpha'$, 并保持三种运算: 对一切 $\alpha, \beta \in \mathfrak{A}$ 和一切 $c \in F$, 有

$$(\alpha+\beta)' = \alpha' + \beta', \quad (c\alpha)' = c\alpha', \quad (\alpha\beta)' = \alpha'\beta' \tag{5}$$

时, 我们称 $\mathfrak{A}$ 与 $\mathfrak{A}'$ 是同构的.

定理 1　*具有单位元素的每个 n 阶线性结合代数与一个 $n \times n$ 矩阵代数同构.*

证明　代数 $\mathfrak{A}$ 是元素 $\boldsymbol{\xi}$ 的向量空间. 同 $\mathfrak{A}$ 中每个元素 $\boldsymbol{\alpha}$ 相联系的变换 T, 是通过右乘 $\mathfrak{A}$ 中任意元素 $\boldsymbol{\xi}$(即 $\boldsymbol{\xi}T = \boldsymbol{\xi\alpha}$) 而得到. 因为如 (4) 中所示的乘法是双线性的, 所以 T 是线性变换. 因为现在单位元素是 1, 由 $1\boldsymbol{\alpha} = 1\boldsymbol{\beta}$ 推出 $\boldsymbol{\alpha} = \boldsymbol{\beta}$, 所以不同的元素 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 导出不同的变换 T 和 U. 此外, 由代数公设得到

$$\boldsymbol{\xi}(\boldsymbol{\alpha}+\boldsymbol{\beta}) = \boldsymbol{\xi\alpha} + \boldsymbol{\xi\beta}, \quad \boldsymbol{\xi}(c\boldsymbol{\alpha}) = c(\boldsymbol{\xi\alpha}), \quad \boldsymbol{\xi}(\boldsymbol{\alpha\beta}) = (\boldsymbol{\xi\alpha})\boldsymbol{\beta},$$

所以相应的变换是 $\boldsymbol{\alpha}+\boldsymbol{\beta} \longmapsto T+U, c\boldsymbol{\alpha} \longmapsto cT, \boldsymbol{\alpha\beta} \longmapsto TU$. 这意味着对应 $\boldsymbol{\alpha} \longmapsto T$ 是给定的代数与 $\mathfrak{A}$ 上线性变换的代数之间的同构. 这些变换又可用矩阵同构地表示出来, 因此定理成立.

习　题

1. (a) 证明: 在任意环中有 $(-a)(-b) = ab, -(-a) = a$.
   (b) 证明: 对所有 $a, a0 = 0a = 0$, 并且单位元素 1 是唯一的.
2. 证明: 由 (2) 定义的直和实际上是一个环.
3. 证明: 两个整环的直和不是整环.
4. 写出 n 个已知环的直和的定义, 并证明它是一个环.
5. 证明: 域 F 上两个线性代数的直和, 在适当定义 "数乘" 运算之后, 就可成为 F 上的线性代数.
6. 证明正文中关于描述子环 S 的命题.
7. 证明: 线性代数的零元素 0, 满足 $\xi \cdot 0 = 0 = 0 \cdot \xi$, 对一切 ξ.
8. "对偶数" 的代数是可除代数吗? 给出证明.
9. 证明下列各系统是线性代数:
   (a) 对所有 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 满足条件 $\boldsymbol{\alpha} \cdot \boldsymbol{\beta} = 0$ 的向量空间 V_n,
   (b) 所有 $n \times n$ 三角形矩阵 (对角线下面的元素都为零).

10. 证明：如果 $\boldsymbol{P}$ 是 F 上任意 $n\times n$ 可逆矩阵，那么 $\boldsymbol{A}\mapsto\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 是 $M_n(F)$ 的一个自同构. 推广这个结果.
11. 证明：同每个 $n\times n$ 矩阵可交换的 $n\times n$ 矩阵 $\boldsymbol{A}$ 一定是标量矩阵. (提示：$\boldsymbol{A}$ 同每个 $\boldsymbol{E}_{ij}$ 可交换.)
12. 设 $\mathscr{A}$ 是一个代数，证明：$\mathscr{A}$ 中所有那些与 $\mathscr{A}$ 的每个元素可交换的元素 z 组成的集合 $\mathscr{I}$ 是 $\mathscr{A}$ 的子代数. (它称为 $\mathscr{A}$ 的中心.)

13.2 同　态

给定两个环 A 和 A' 以及对应 $a\longmapsto aH$，如果对 A 中每个元素 a，aH 是 A' 中唯一确定的元素，并且对 A 中一切元素 a,b，有

$$(a+b)H=aH+bH,\quad (ab)H=(aH)(bH),\quad 1H=1', \tag{6}$$

则我们称对应 $a\longmapsto aH$ 是 A 到 A' 的同态. 简单地说，正如 3.3 节中交换环的情形一样，同态是一个保持单位元素、保持和与积的映射. 同群的情形一样，映上的同态也称为满同态.

从环 A 到 A' 的同态 H 一定是由 A 的加法群到 A' 的加法群的同态，所以 H 具有 6.11 节中对于群已经证明了的性质

$$0H=0',\quad (-a)H=-(aH),\quad (a-b)H=aH-bH, \tag{7}$$

这里 $0'$ 是环 A' 的零元素，即 A' 的加法群的单位元素.

我们所熟悉的把每个整数 a 映射到模 m 剩余类的对应 $a\longmapsto a_m$ 是整数环 $\mathbf{Z}$ 到 $\mathbf{Z}_m$ 的同态. 如果 $f(x)$ 是系数在整环 D 中的任意多项式，那么用 D 中固定元素 b "替换" $f(x)$ 中的 x 而得到的对应 $f(x)\longmapsto f(b)$ 是多项式整环 $D[x]$ 到 D 的一个同态，因为未定元 x 的多项式形式的加法法则和乘法法则，对于相应的 b 的多项式表达式，当然适用. 如果 $\mathbf{Q}[x]$ 是有理系数的多项式环，那么对应 $f(x)\longmapsto f(\sqrt{2})$ 是多项式环 $\mathbf{Q}[x]$ 到由所有数 $a+b\sqrt{2}$ 构成的域上的满同态 (见 2.1 节的讨论). 两个环 A 与 B 的直和 $A\oplus B$，通过对应 $(a,b)\longmapsto b$ 满同态地映到被加项 B；正好根据直和运算的定义 (2)，这个对应保持和与积.

为了明确地描述一个具体的同态，我们自然要问，什么时候第一个环中两个元素 a 和 b 在第二个环中有相同的像. 根据法则 (7)，只有当它们的差的像 $(a-b)H=0'$ 时，才可能发生这种情况. 因此我们寻求这样元素的集合，这些元素通过 H 映射到 A' 的零元素 $0'$ 上. 例如，同态 $\mathbf{Z}\longrightarrow\mathbf{Z}_m$ 把模 m 的所有倍数 km 映上到零. 所有这些倍数的集合在减法之下是封闭的，还有，这个集合的元素同 $\mathbf{Z}$ 的任何元素相乘后仍在这个集合中. 类似地，同态 $f(x)\longmapsto f(b)$ 把所有可被 $x-b$ 整除的多项式映射

到零, 再没有其他多项式可以映射到零. 所有这些多项式组成的集合 S 在减法和用 $D[x]$ 的所有元素 (不管这些元素是否在 S 中) 与它相乘的运算之下是封闭的. 这两个例子蕴含着下面的定义和定理 (参看 3.8 节).

定义 环 A 中的理想C 是具有下述性质的 A 中非空集合:

(i) c_1 和 c_2 在 C 中, 可推出 $c_1 - c_2$ 在 C 中;

(ii) c 在 C 中, a 在 A 中, 可推出 ac 和 ca 都在 C 中.

定理 2 在环 A 的任意同态 H 中, 由所有映射到零的元素组成的集合是 A 的一个理想.

为了在一般情形下证明定理 2, 设 C 是 A 中所有满足 $cH = 0'$ 的元素 c 的集合, 其中 $0'$ 是像 A' 的零元素. 那么, 对 A 中的任意元素 a, 有 $(ac)H = (aH)(cH) = (aH)0' = 0'$ 和 $(ca)H = (cH)(aH) = 0'$, 这就证明了性质 (ii). 此外, 根据 (7), 由 $c_1H = c_2H = 0'$, 可得到

$$(c_1 - c_2)H = c_1H - c_2H = 0' - 0' = 0',$$

因此证明了性质 (i).

这个结果表明, 环中的理想类似于群中的正规子群, 为表示这种类似, 我们称通过同态 H 映射到零的所有元素组成的集合为 H 的核. 当 H 是满射 (即满同态) 时, 我们称环 B 是环 A 在同态 H 之下的满同态像, 所以每个元素 $b \in B$ 是某个 $a \in A$ 在 H 之下的像 aH.

定理 3 环 A 的满同态像由它的核确定 (除同构外).

证明 我们必须证明, 如果 H 和 K 分别是 A 到 A' 和 A'' 上的满同态, 并且 $aH = 0'$ 当且仅当 $aK = 0''$, 那么 A' 和 A'' 是同构的. 很自然地, 设元素 $a' \in A'$ 对应于 $a'' \in A''$ 当且仅当这两个元素在 A 中有公共的像源 a, 所以对某个 a, 当 $aH = a', aK = a''$ 时, 有 $a' \longleftrightarrow a''$. 这个对应是一一的: 在这个对应下, 对 A' 中每个 a', 在 A'' 中有一个且只有一个 a'' 与 a' 对应. 为证明这一点, 首先注意, 对 A' 中每个 a', 在 A 中至少有一个像源 a, 因此在 A'' 中至少有一个 $a'' = aK$ 与 a' 对应. 其次, 如果 $a' \longleftrightarrow a'', a' \longleftrightarrow b''$, 那么对 A 中某两个 a, b, 有

$$aH = a', \quad aK = a'', \quad bH = a', \quad bK = b'',$$

因此 $(a - b)H = a' - a' = 0'$, 根据假设可推出 $0'' = (a - b)K = a'' - b''$. 这个对应也保持和与积, 这是因为如果 $a' \longleftrightarrow a'', b' \longleftrightarrow b''$ 那么

$$a' + b' = (a + b)H \longleftrightarrow (a + b)K = a'' + b'',$$
$$a'b' = (ab)H \longleftrightarrow (ab)K = a''b'',$$

这里 a 是 a' 和 a'' 的公共像源, b 是 b' 和 b'' 的公共像源.

理想的两个性质 (i) 和 (ii) 有几个直接推论. 任意理想 C 包含某元素 c, 因此 (i) 表明 $c-c=0$ 在 C 中, 所以对 C 中任意 c, $0-c=-c$ 也在 C 中. 根据性质 (i), 我们得到, C 中任意两个元素之和 $c_1+c_2=c_1-(-c_2)$ 也在 C 中. 于是, 因为 $1\in A$, 所以 A 的非空子集 C 是 A 的理想当且仅当任意线性组合 $a_1c_1\pm a_2c_2$ 和 $c_1a_1\pm c_2a_2$ 在 C 中, 其中 c_1 和 c_2 在 C 中, a_1 和 a_2 在 A 中. 特别是, A 的理想不一定是 A 的子环, 因为它可以不包含 A 的单位元素. 整个环 A 和仅由一个元素 0 组成的 $\{0\}$ 总是任意环 A 的理想. 它们称为环 A 的*假理想*. 任意其他理想称为*真理想*. 相应地, 环 A 的*真满同态*是这样一个同态, 它的核是 A 的真理想, 所以真满同态不是一个同构 (同构只把 $\{0\}$ 映到 $0'$).

定理 4　*可除环没有真满同态像.*

证明　只须证明可除环 D 没有真理想. 设 C 是 D 中任意理想, 它不是理想 $\{0\}$, 于是它包含一个元素 $c\neq 0$. 由性质 (ii), C 包含 $1=c^{-1}c$, 再由性质 (ii), C 包含整个可除环的任意元素 $a=a\cdot 1$. 因此 C 不是真理想, 如断言所述.　证毕

如果 b 是*交换环*A 的元素, 那么 b 的所有倍数 xb(其中变量 $x\in A$) 的集合 (b) 是一个理想, 因为可以验证性质 (i) 和 (ii). 这个理想 (b) 称为*主理想*, 它是 A 中包含 b 的最小理想. 我们回忆一下, 由 1.7 节定理 6 可知, 整数环 $\mathbf{Z}$ 中的每个理想都是主理想. 根据 3.8 节定理 11, 在任意域 F 上的一个未定元的多项式整环 $F[x]$ 中, 上述结论同样成立.

在两上变量的有理系数多项式环 $\mathbf{Q}[x,y]$ 中, 常数项为零的所有多项式的集合 C 是一个理想. 它并不是主理想, 因为两个多项式 x 和 y 虽然都在 C 中, 但它们不能都是其中一个多项式的倍式, 也不能同时是同一个多项式 $f(x,y)$ 的倍式. 虽然这个理想 C 不能由任意单个多项式 $f(x,y)$ 生成, 但是它的所有元素都可用含有多项式系数的线性组合 $xg(x,y)+yh(x,y)$ 来表示, 所以整个理想是通过两个生成元 x 和 y 的线性组合给出.

现在考虑由交换环 A 中任意给定的有限个元素的集合生成的理想. 如果一个理想 C 包含元素 $c_1,c_2,\cdots,c_m$, 那么它必包含所有线性组合 $\sum_i x_ic_i$, 其中系数 x_i 在 A 中. 而集合

$$(c_1,c_2,\cdots,c_m)=\left[\text{所有元素 }\sum_i x_ic_i,\quad x_i\in A\right] \tag{8}$$

本身是一个理想, 这是因为

$$\sum_i x_ic_i-\sum_i y_ic_i=\sum_i(x_i-y_i)c_i,\quad a\left(\sum_i x_ic_i\right)=\sum_i(ax_i)c_i,$$

也就是说, 这个集合具有关于理想所要求的性质 (i) 和 (ii). 因为 A 具有单位元素 1,

所以每个元素 c_i 必然是集合 (8) 中的一个元素

$$c_i = 0 \cdot c_1 + \cdots + 0 \cdot c_{i-1} + 1 \cdot c_i + 0 \cdot c_{i+1} + \cdots + 0 \cdot c_m.$$

因此, 由 (8) 式定义的集合 $(c_1, \cdots, c_m)$ 是包含 $c_1, \cdots, c_m$ 的 A 的一个理想, 并且它还被包含在包含所有 c_i 的任何理想中, 它称为以 $c_1, \cdots, c_m$ 为基的理想. (这种基元素不能同向量空间的基相类比, 因为 $x_1c_1 + \cdots + x_mc_m = 0$ 不一定能推出 $c_1 = \cdots = c_m = 0$.)

在大多数熟悉的整环中, 每个理想都有一组有限基, 但是也存在一些整环情况并非如此.

习 题

1. 在下列映射中, 哪些是同态, 为什么? 对于是同态的映射, 描述映射到零的理想.
   (a) $a \mapsto 5a$, a 是 $\mathbf{Z}$ 中整数.
   (b) $f(x) \mapsto f(\omega)$, $f(x)$ 是 $\mathbf{Q}[x]$ 中多项式, ω 是三次单位根.
   (c) $f(x,y) \mapsto f(t,t)$, 是 $F[x,y]$ 到 $F[t]$ 的映射 (x, y, t 是未定元).
2. 证明: 交换环的每个同态像是可交换的.
3. 在多项式 $f(x,y) = a + b_1x + b_2y + c_1x^2 + c_2xy + c_3y^2 + \cdots$ 组成的环 $\mathbf{Q}[x,y]$ 中, 下列多项式集合中哪些是理想? 如果那个集合是理想, 就求出它的基.
   (a) 常数项为零 ($a = 0$) 的所有多项式 $f(x,y)$.
   (b) 不含 x 的所有多项式 $f(x,y)(b_1 = c_1 = c_2 = \cdots = 0)$.
   (c) 不含二次项的所有多项式 ($c_1 = c_2 = c_3 = 0$).
4. (a) 求出 $\mathbf{Z}_6$ 中的所有理想. (b) 求出 $\mathbf{Z}_6$ 的所有同态像.
5. 详细证明: 第 1 章中定义的环 $\mathbf{Z}/(m) = \mathbf{Z}_m$ 是 $\mathbf{Z}$ 的唯一的真满同态像.
6. (a) 对每个 m 求出 $\mathbf{Z}_m$ 中的所有理想. (b) 求出 $\mathbf{Z}_m$ 的所有满同态像.
7. 求出两个域的直和中的所有理想. 并推广这个结果.
8. 求出直和 $\mathbf{Z} \oplus \mathbf{Z}$ 中的所有理想, 这里 $\mathbf{Z}$ 是整数环.

*9. 证明: 如果 C_1 和 C_2 分别是环 A_1 和 A_2 中的理想, 那么直和 $C_1 \oplus C_2$ 是直和 $A_1 \oplus A_2$ 中的理想, 并且这个直和中的每个理想都具有这种形式.

10. 证明: 在整环中, $(a) = (b)$ 当且仅当 a 和 b 是相伴的 (3.6 节).
11. 证明: 如果 A 是一个交换环, 其中每个理想都是主理想, 那么 A 中任意两个元素 a 和 b 都有最大公因子 d, 并可表示成 $d = ra + sb$.

*12. 设 A 是包含域 F 的环, A 和 F 具有同一个单位元素 (例如 A 可以是域 F 上的多项式环). 证明: A 的每个真同态像包含一个与 F 同构的子域.

*13. 设 $\mathbf{Z}_p^*$ 是由所有有理数 $\dfrac{m}{n}$(其中分母与给定的素数 p 互素) 组成的环. 证明: $\mathbf{Z}_p^*$ 中每个真理想具有形式 (p^k), 其中 k 是某正整数.

13.3 商 环

对环的每个同态, 都存在对应的理想, 该理想的元素在同态下映射到零. 反过来, 给定一个理想, 我们现在将构造一个相应的同态像. 环 A 中的理想 C 是 A 的加法群的子群. A 中每个元素 a 属于一个陪集, 这个陪集常常称为剩余类 $a' = a + C$, 它是由所有和 $a + c$(变量 $c \in C$) 组成. 两个元素 a_1 和 a_2 属于同一个陪集当且仅当它们的差在这个理想 C 中. 因为加法是可交换的, 所以 C 是加法群 A 的正规子群, 因此 C 的所有陪集构成阿贝耳商群, 在这个商群中, 两个陪集的和是另一个陪集, 它是通过把两个代表元相加而得到, 即

$$(a_1 + C) + (a_2 + C) = (a_1 + a_2) + C. \tag{9}$$

6.13 节中已经证明, 这个和不依赖于在给定的陪集中元素 a_1 和 a_2 的选择.

为了构造两个陪集的乘积, 在第一个陪集中选取任意元素 $a_1 + c_1$, 在第二个陪集中选取任意元素 $a_2 + c_2$. 乘积

$$(a_1 + c_1)(a_2 + c_2) = a_1a_2 + (a_1c_2 + c_1a_2 + c_1c_2) = a_1a_2 + c'$$

总是陪集 $a_1a_2 + C$ 中的一个元素, 因为根据理想的性质 (ii), a_1c_2, c_1a_2, c_1c_2 这些项都在理想 C 中. 因此, 第一个陪集中的元素与第二个陪集中的元素的所有乘积都在同一个陪集中, 这个乘积陪集是

$$(a_1 + C)(a_2 + C) = a_1a_2 + C. \tag{10}$$

陪集乘法的结合律和分配律立即从 A 中相应的定律可以得到. 包含 1 的陪集起单位元素的作用, 所以 A 中 C 的全体陪集构成一个环.

正是根据陪集运算的定义 (9) 和 (10), 把 A 的每个元素映射到它的陪集的对应 $a \longmapsto a' = a + C$ 是一个满同态. 在这个满同态像中, 零元素是陪集 $0 + C$, 所以 C 的元素都被映射到零. 这些结果可以总结如下:

定理 5 *在定义 (9) 和 (10) 之下, 环 A 中任意理想 C 的全体陪集构成一个环, 称为商环*[①] *A/C. 函数 $a \longmapsto a + C$ 把 A 的每个元素映射到包含它的陪集, 它是 A 到商环 A/C 的一个满同态, 而且这个满同态核是给定的理想 C.*

推论 1 *如果 A 是可交换的, 那么 A/C 也是可交换的.*

理想与同态的关系现在已经齐全了. 特别是, 定理 3 关于唯一性的断言可以改述如下:

① 环 A/C 也常被称为剩余类环, 因为它的元素是 C 在 A 中的剩余类 (陪集).

推论 2 *如果满同态 H 把 A 映射到 A', 并有同态核 C, 那么 A' 与商环 A/C 同构.*

模 m 整数的环 $\mathbf{Z}_m$ 现在可以描述成商环 $\mathbf{Z}/(m)$. 反过来, 由这个例子所启发, 当 $(a-b)\in C$ 时, 我们常常写成 $a\equiv b(C)$, 并且说 a 和 b 是同余的模环 R 的理想 C.

商环的每个性质都反映在它的生成理想 C 的相应的性质中. 为了解释这个原理, 我们定义极大理想和素理想. 如果 A 中包含理想 C 的理想只能是 C 和 A 本身, 则我们称 $C<A$ 是*极大理想*[①]. 如果 A 中理想 P 包含乘积 ab 时, 至少包含其中一个因子 a 或 b, 则我们称 P 为*素理想*.

在交换环中, 素理想起着特殊的作用. 例如, 在整数环 $\mathbf{Z}$ 中, 主理想 (p) 是素理想当且仅当 p 是素数, 因为当 p 是素数, 而不是其他情形时, 两个整数的乘积 ab 是 p 的倍数当且仅当有一个因子是 p 的倍数.

定理 6 *如果 A 是交换环, 那么商环 A/C 是整环当且仅当 C 是素理想, 商环 A/C 是域当且仅当 C 是 A 的极大理想.*

证明 交换环 A/C 是整环当且仅当它没有零因子 (1.2 节定理 1). 这个条件用公式写成

$$a'b'=0,\ \text{仅当}\ a'=0\ \text{或}\ b'=0, \tag{11}$$

这里 a' 和 b' 分别是元素 a 和 b 在 A 中的陪集. 现在 C 的陪集 a' 是零当且仅当 a 在理想 C 中, 则上述条件可以改写成

$$ab\ \text{在}\ C\text{中, 仅当}\ a\ \text{在}\ C\ \text{中或}\ b\ \text{在}\ C\ \text{中}. \tag{12}$$

这恰好就是素理想 C 的定义.

其次, 假定 C 是极大理想, 并设 b 是 A 中不属于 C 的任意元素. 那么可以证明, 所有元素 $c+bx$(其中任意 $c\in C$, 任意 $x\in A$) 组成的集合是一个理想. 这个理想包含 C, 并包含不在 C 中的元素 b. 因为 C 是极大理想, 所以这个理想一定是整个环 A. 特别是, 单位元素 1 在这个理想中, 所以对某个 a, 有 $1=c+ba$. 用陪集来表示, 这个方程写成 $1'=b'a'$. 于是, 对任意陪集 $b'=b+C\neq C$, 我们已求出互逆陪集 $a'=a+C$, 这就是说, 陪集组成的交换环是一个域. 反过来, 如果 A/C 是一个域, 我们可以证明 C 是极大理想 (习题 10). 证毕

因为每个域是一个整环, 所以定理 6 意味着每个极大理想是素理想. 然而反过来, 素理想不一定是极大理想. 例如, 考虑同态 $f(x,y)\longmapsto f(0,y)$, 它把系数在域 F 上的未定元 x 和 y 的多项式整环 $F[x,y]$ 映射到较小的整环 $F[y]$ 上. 因此映射到零的理想是主理想 (x), 它是由 x 的所有倍式 (多项式) 组成. 因为像环 $F[y]$ 实际上是一个整环, 所以这个理想 (x) 是一个素理想, 这也可以直接验证. 但是 $F[y]$ 不是域,

① "极大理想" 有时用 "无因子理想" 代替.

所以 (x) 不可能是极大理想. 实际上, 它包含在较大的理想 (x,y) 之中, (x,y) 是由常数项为零的所有多项式组成.

习　题

1. 证明：陪集的乘法满足结合律和分配律.
2. 设模一个理想 $C \leqslant A$ 的同余定义为：$a \equiv b(\mathrm{mod} C)$ 当且仅当 $a-b$ 在 C 中. 证明：同余可以相加和相乘, C 的陪集是由相互同余的元素组成.
3. 详细证明定理 5 的推论 1.
4. 求出整数环 $\mathbf{Z}$ 中所有素理想.
5. 求出域 F 上多项式环 $F[x]$ 中的所有素理想和所有极大理想.

*6. 不用定理 6 证明：整环中每个极大理想都是素理想.

*7. 在整系数多项式整环 $\mathbf{Z}[x]$ 中, 求出不是极大理想的素理想.

8. 证明：在所有数 $a+b\omega(a,b$ 为整数, ω 是虚三次单位根) 组成的整环 $\mathbf{Z}[\omega]$ 中, (2) 是素理想. 并描述商环 $\mathbf{Z}[\omega]/(2)$.
9. 在多项式环 $\mathbf{Q}[x,y]$ 中, 下列理想中哪些是素理想? 哪些是极大理想?

   (a) (x^2),　(b) $(x-2,y-3)$,　(c) $(y-3)$,

   (d) (x^2+1),　(e) (x^2-1),　(f) $(x^2+1,y-3)$.
10. 证明：如果商环 A/C 是一个域, 那么 C 是极大理想.
11. 求出一个熟悉的环与下面每个商环 A/C 同构：

    (a) $A=\mathbf{Q}[x], C=(x-2)$;　(b) $A=\mathbf{Q}[x], C=(x^2+1)$;

    (c) $A=\mathbf{Q}[x,y],\ C=(x,y-1)$;　(d) $A=\mathbf{Z}[x], C=(3,x)$;

    (e) $A=\mathbf{Z}_p^*,\ C=(p)$, 这里 $\mathbf{Z}_p^*$ 的定义见 13.2 节的习题 13.
12. (第二同构定理) 设 $C>D$ 是环 A 中的两个理想.

    (a) 证明：商 C/D 是 A/D 中的理想.

    (b) 证明：A/C 与 $(A/D)/(C/D)$ 同构. (提示：两个同态的乘积是一个同态.)

*13.4 理想的代数

理想之间的包含同数之间的整除性有着密切的关系. 在整数环 $\mathbf{Z}$ 中, $n|m$ 意味着 $m=an$, 因此 m 的每个倍数是 n 的倍数. n 的倍数组成主理想 (n), 所以条件 $n|m$ 意味着 (m) 包含在 (n) 中. 反过来, $(m)\subset(n)$ 特别意味着 m 在 (n) 中, 因此 $m=an$. 所以

$$(m)\subset(n)\quad \text{当且仅当}\quad n|m.$$

更一般地，在任意交换环 R 中，$(b) \subset (a)$ 可推出，对某个 $x \in R$ 有 $b = ax$，即 $a|b$. 反过来，如果 $a|b$，那么对某个 $x \in R$ 有 $b = ax$，于是对所有的 $by \in (b)$，有 $by = axy \in (a)$，因此 $(b) \subset (a)$. 这就证明了

定理 7 *在任意交换环 R 中，*

$$(b) \subset (a) \quad \text{当且仅当} \quad a|b. \tag{13}$$

但要注意："较大"的数对应着"较小"的理想；例如，6 的所有倍数组成的理想 (6) 真正地包含在所有偶数组成的理想 (2) 中.

最大公因子和最小公倍数在理想理论中也有相应的解释. 整数 n 和 k 的最小公倍数 m 是 n 和 k 的倍数，并且是 n 和 k 的其他每个公倍数的因子. 于是，m 的所有倍数的集合 (m) 是 n 和 k 的所有公倍数的集合，刚好也是主理想 (n) 和 (k) 的公共元素组成的集合. 这个情况可以推广到任意环（不一定是交换环）的任意理想上，如下所述.

可以证明，环 A 的任意两个理想 B 和 C 的交 $B \cap C$ 是一个理想. 设 D 是 A 的任意其他理想，则理想 $B \cap C$ 具有三个性质：

(i) $B \cap C \subset B$, (ii) $B \cap C \subset C$,

(iii) 由 $D \subset B$ 和 $D \subset C$ 可推出 $D \subset B \cap C$.

于是在格论意义之下，这个交是 B 和 C 的最大下界.

对偶于交的是两个理想的*和*. 如果 B 与 C 是 A 中两个理想，我们则可验证集合

$$B + C = [\text{所有的和 } b + c, \text{其中 } b \in B, c \in C] \tag{14}$$

是 A 中一个理想. 因为包含 B 和 C 的任意理想一定包含所有的和 $b + c$，所以这个理想 $B + C$ 包含 B 和 C，并且包含在每个包含 B 和 C 的理想中. 于是在格论意义下 $B + C$ 是 B 与 C 的最小上界也是 B 与 C 的并.

定理 8 *由 (14) 式的和 $B + C$ 给出的并与由 $B \cap C$ 给出的交，在通常包含关系之下，环 A 中的全体理想构成一个格.*

如果整数 m 和 n 有最大公因子 d，那么理想之和 $(m) + (n)$ 恰好是主理想 (d). 这是因为根据 (13)，有 $(d) \supset (m)$ 和 $(d) \supset (n)$；由于 d 有表达式 $d = rm + sn$，所以包含 m 和 n 的任意理想一定包含 d，因而也包含 (d) 的所有元素. 因此，(d) 是 (m) 和 (n) 的并，即 $(d) = (m) + (n)$.

前面的研究可以推广如下：

引理 *在交换环 R 中，两个主理想的和 $(b) + (c)$ 本身是主理想 (d) 当且仅当 d 是 b 和 c 的最大公因子.*

我们把引理的证明留给读者.

一般地, 在交换环中, 如果理想 B 和 C 是由如下的基生成的:

$$B = (b_1, \cdots, b_m), \quad C = (c_1, \cdots, c_n), \tag{15}$$

那么我们有, 对任意 $b + c \in B + C$, 按 (8) 式得

$$b + c = \sum_i x_i b_i + \sum_j y_j c_j,$$

也就是说, $B + C$ 是由 $b_1, \cdots, b_m$ 和 $c_1, \cdots, c_n$ 生成, 所以

$$(b_1, \cdots, b_m) + (c_1, \cdots c_n) = (b_1, \cdots, b_m, c_1, \cdots, c_n). \tag{16}$$

这个法则同基的自然变换结合起来, 可以用来明显地计算出整数的最大公因子. 例如,

$$\begin{aligned}(336) + (270) &= (336, 270) = (336 - 270, 270) = (66, 270)\\ &= (66, 270 - 4 \times 66) = (66, 6) = (6),\end{aligned}$$

所以 336 和 270 的最大公因子是 6.

在任意交换环中, 我们还可以定义任意两个理想 B 与 C 的乘积BC,

$$BC = [\text{所有和 } b_1c_1 + \cdots + b_mc_m, \text{其中 } b_i \in B, c_j \in C]. \tag{17}$$

实际上, 这个集合是一个理想, 它是由所有的乘积 bc 生成的, 其中因子 b 在 B 中, 因子 c 在 C 中, 所以它也是包含所有这些乘积的最小理想. 特别是, 两个主理想 (b) 和 (c) 的乘积就是由已知元素 b 和 c 的乘积生成的主理想 (bc). 更一般地, 如果理想 B 和 C 是由 (15) 式的基确定的, 任意乘积 bc 具有形式

$$bc = \left(\sum_i x_i b_i\right)\left(\sum_j y_j c_j\right) = \sum_{i,j} (x_i y_j)(b_i c_j),$$

因此乘积理想 BC 有基

$$BC = (b_1c_1, b_1c_2, \cdots, b_mc_{n-1}, b_mc_n). \tag{18}$$

这种乘积对于代数数论 (14.10 节) 是有用的.

习 题

1. 详细证明: $B \cap C$ 与 $B + C$ 总是理想.

2. 证明：(17) 式的乘积 BC 是一个理想.
3. 画出 $\mathbf{Z}_{24}$ 中所有理想的格图.
4. 设 $f(x)$ 和 $g(x)$ 是域上的多项式, $d(x)$ 是它们的最大公因式. 证明：$(f(x))+(g(x))=(d(x))$.
5. 用理想基的方法计算最大公因子 (280, 396) 和 (8624, 12825).
6. 证明：整数环 $\mathbf{Z}$ 中每个理想可唯一地表示成素理想的乘积.
7. 证明下列变换理想的基的法则：

$$(c_1, c_2, \cdots, c_m) = (c_1 + xc_2, c_2, \cdots, c_m),$$
$$(xc_1, c_2, \cdots, c_m) = (c_1, c_2, \cdots, c_m).$$

8. 在 $R[x,y]$ 中, 化简下面理想的基：

$$(x^2+y,\ 3y,\ 4x^3+x^2), \quad (x^2+3xy+y^2,\ 2x^2-y^2, x^2+6xy,\ x^3+y^2).$$

9. (a) 证明：在任意交换环中, $BC \subset B \cap C$.
   (b) 给出一个例子说明 $BC < B \cap C$ 是可能的.
   (c) 证明：$B(C+D) = BC + BD$.

*10. 证明：任意环中所有理想构成的格是一个模 (按 11.7 节习题 10 的定义).

11. 在交换环 A 中, 设 $B:C$ 表示满足 $xc \in B(c \in C)$ 的所有元素 x 组成的集合.
   (a) 证明：如果 B 和 C 是理想, 那么 $B:C$ 也是 A 中的理想 (称为"理想商").
   (b) 证明：$(B_1 \cap B_2):C = (B_1:C) \cap (B_2:C)$.
   (c) 证明：$B:C$ 是满足 $CX \subset B$ 的所有理想 X 的最小上界 (并).
12. 证明：如果环 R 包含理想 B 和 C, 并满足 $B \cap C = 0$, $B + C = R$, 那么 R 同构于 B 与 C 的直和.

13.5 多项式理想

理想的概念在现代代数几何中是基本的. 当我们研究三维空间的代数曲线时, 由于引入理想的概念, 使推理立即变得明显了.

一般在 n 维向量空间 F^n 中, 一个 (仿射)代数簇定义为点集合 V, 这个集合中的所有点 $(x_1, \cdots, x_n)$ 都满足有限个适当的多项式方程的方程组

$$f_1(x_1, \cdots, x_n) = 0, \cdots, f_m(x_1, \cdots, x_n) = 0. \tag{19}$$

例如, 在 $\mathbf{R}^3$ 中, 平行于 xy 平面且在 xy 平面上方两个单位的平面中的圆 C(圆心在 z 轴上, 半径为 2) 通常可解析地描述为空间中满足联立方程

$$x^2 + y^2 - 4 = 0, \quad z - 2 = 0 \tag{20}$$

的点 (x,y,z) 的集合. 这些方程把曲线 C 描绘成圆柱面与一平面的交线. 但是也可以用等价的联立方程

$$x^2+y^2+z^2-8=0, \quad z-2=0, \tag{21}$$

以同样的精确度把 C 描绘成一个球面与平面 $z=2$ 的交线. 还可能有其他的描述, 例如用方程组

$$x^2+y^2-4=0, \quad x^2+y^2-2z=0. \tag{22}$$

这些方程把 C 描绘成圆柱面与旋转抛物面 $x^2+y^2=2z$ 的交线.

我们可以通过用全体这样的多项式方程 (即曲线 C 的点所适合的方程) 描绘 C 来避免上述的含糊. 如果 $f(x,y,z)$ 和 $g(x,y,z)$ 是任意两个多项式, 它们的值在 C 上恒等于零, 那么它们的和与差在 C 上也恒等于零. 同样地, 不管什么样的多项式 $a(x,y,z)$ 与 $f(x,y,z)$ 的乘积 $a(x,y,z)f(x,y,z)$ 在 C 上也恒等于零. 这就意味着, *在 C 上的值恒等于零的所有多项式组成的集合是一个理想*. 那么这个理想 (而不是它的任意一对个别的元素) 就是 C 的根本的描述. 我们现在将证明所有这样的方程组成的集合是一个理想.

定理 9 *在 F^n 中, 在给定的集合 S 上恒等于零的所有多项式的集合 $J(S)$ 是 $F[x_1,\cdots,x_n]$ 中的一个理想.*

因为如果 $p(x_1,\cdots,x_n)$ 在给定的点上等于零, 那么 p 的所有倍式在该点上也等于零, 而当 p 和 q 在给定的点上等于零时, 则 $p\pm q$ 在该点上也等于零. 对于在给定的集合 S 上恒等于零的多项式, 上述结论同样正确. 实际上, $J(S)$ 刚好是在不同的点 $\xi\in S$ 上等于零的多项式理想 $J(\xi)$ 的交.

例如, 上面所讨论的圆 C 的情形中, $J(C)$ 是由所有线性组合

$$h(x,y,z)=a(x,y,z)(x^2+y^2-4)+b(x,y,z)(z-2) \tag{23}$$

构成的理想, 其中系数为多项式 $a(x,y,z)$ 和 $b(x,y,z)$. 这就是说, $J(C)$ 就是以 x^2+y^2-4 和 $z-2$ 为基的理想 $(x^2+y^2-4,z-2)$. (21) 中的多项式生成同一个理想, 因为这些多项式是 (20) 的那两个多项式的线性组合, 而反过来, (20) 的多项式可以由 (21) 的两个多项式的线性组合而得到. 于是由这条曲线确定的多项式理想有各种各样的基.

$$(x^2+y^2-4,z-2)=(x^2+y^2+z^2-8,z-2)=(x^2+y^2-2z,z-2). \tag{24}$$

商环 $\mathbf{R}[x,y,z]/(x^2+y^2-4,z-2)$ 有一个重要意义. 即它与定义在 C 上的所有函数组成的环同构 (参看 3.2 节), 这些函数可定义为变量 x,y,z 的多项式. 显然, 它与 $\mathbf{R}[x,y]/(x^2+y^2-1)$ 同构, 因此用通常的同化规则可知, 它与所有三角多项

式$p(\cos\theta, \sin\theta)$ 构成的环同构. 这个商环称为 C 上多项式函数环, 由它扩张成的域称为 C 上有理函数域.

三次挠线$C_3: x = t, y = t^2, z = t^3$ 是一条代数曲线, 它 (不像 C 那样) 可以用参数 t 的多项式函数来定义. 显然, 给定的点 (x, y, z) 在 C_3 上当且仅当 $y = x^2, z = x^3$. 因此 C_3 是 $\mathbf{R}^3$ 中由理想 $M = (y - x^2, z - x^3)$ 定义的代数曲线.

根据定义, 多项式$p(x, y, z)$ 在 C_3 上恒等于零当且仅当对所有 $t \in \mathbf{R}, p(t, t^2, t^3) = 0$. 现在考虑同态[①]

$$f(x, y, z) \longmapsto f(t, t^2, t^3) \quad (t \text{ 是未定元}). \tag{25}$$

显然, 对 C_3 上的所有点, 有 $y = x^2, z = x^3$, 这表明 $y - x^2$ 和 $z - x^3$ 将位于我们的理想 M 之中. 但是反过来, 我们注意, 变量替换 $y = y' + x^2, z = z' + x^3$ 将把任意多项式 $f(x, y, z)$ 变为 $f'(x, y', z')$, 并注意, 按照这种形式, 同态 (25) 是

$$f'(x, y', z') \longmapsto f'(t, 0, 0). \tag{25$'$}$$

这个对应把 f' 中包含 y' 或 z' 的项映射到零, 而不是别的, 所以被映射到零的多项式就是线性组合 $g(x, y, z)y' + h(x, y, z)z'$. 因此, 我们的理想 M 恰是以 $y' = y - x^2, z' = z - x^3$ 为基的理想 $(y', z') = (y - x^2, z - x^3)$. 这就把曲线 C_3 表示成一个抛物柱面与另一个柱面的交线. 在 C_3 的进一步分析中, 商环 $\mathbf{R}[x, y, z]/M$ 起着重要作用. 映射 (25) 表明这个商环与多项式环 $\mathbf{R}[t]$ 同构.

两个理想的和有简单的几何解释. 例如, 在 $\mathbf{R}[x, y, z]$ 中, 主理想 $(z - 2)$ 表示平面 $z = 2$, 因为这个理想中所有的多项式 $f(x, y, z)(z - 2)$, 当其中的 x, y, z 用平面 $z = 2$ 上点的坐标代替时, 都恒等于零. 类似地, 主理想 $(x^2 + y^2 - 4)$ 定义了一个以 z 轴为轴的半径为 2 的圆柱面. 根据法则 (16), 这两个理想的和是 $(x^2 + y^2 - 4, z - 2)$. 我们刚刚看到, 这个和 (23) 表示一个圆, 它是平面和圆柱面的交线. 事实上显然有: 两个理想的和所对应的轨迹是各理想所确定的轨迹的交.

反过来, 多项式环 $\mathbf{R}[x_1, \cdots, x_n]$ 中的任意理想 J 确定一个相应的轨迹, 这个轨迹是由 n 维空间中使得对每个多项式 $f \in J$ 有 $f(a_1, \cdots, a_n) = 0$ 的所有的点 $(a_1, \cdots, a_n)$ 组成. 希尔伯特基定理断言: J 具有有限基 $f_1, \cdots, f_m$, 所以相应的轨迹 V 的确是一个代数簇. 可是, 这个簇的理想 $J(V)$ 可以大于给定的理想 J(参看下面的习题 3).

习 题

1. 求 $\mathbf{R}^3$ 中具有参数方程 $x = t + 1, y = t^3, z = t^4 + t^2$ 的曲线所对应的理想.

① 注意, (25) 式定义一个同态这一事实并不显然, 证明它需要把 3.1 节定理 1 的推广.

2. 证明：由两个线性无关的线性多项式生成的任意理想 $(ax+by+cz,\ a'x+b'y+c'z)$ 确定 $\mathbf{R}^3$ 中的一条直线.
3. (a) 证明：在 $\mathbf{R}[x,y,z]$ 中的理想 (x,y) 与 (x^2,xy,y^2) 确定相同的代数簇.
   (b) 证明：任意理想和它的平方确定相同的轨迹.
4. 详细证明：在任意轨迹 C 上恒等于零的 $\mathbf{R}[x_1,\cdots,x_n]$ 中多项式的集合是一个理想.
5. (a) 在三维空间中, $xy=0$ 所确定的轨迹是什么?
   (b) 证明：由两个主理想的*乘积*所确定的轨迹是各主理想所确定的轨迹的*并*.
   (c) 把 (b) 的结果推广到任意理想上. (*提示*：如果在乘积的轨迹中有一点不在第一个因子所确定的轨迹上, 那么在第一个理想中至少有一个多项式在这点不为零.)
   (d) 两个理想的交所确定的轨迹是什么?
6. (a) 计算下面"双有理"变换的逆：

$$T: x'=x,\quad y'=y-x^2,\quad z'=z+y+x^3$$

   (b) 证明：所有形为 $x'=x, y'=y+p(x),\quad z=z'+q(x,y)(p,q$ 是多项式) 的变量替换组成的集合是一个群.
   (c) 证明：每个这样的替换诱导出一个环 $\mathbf{R}[x,y,z]$ 上的*自同构*.
7. (a) 证明：如果 H 是交换环 A 中一个理想, H 的*根式*是由 A 中所有使得某个幂 $x^m\in H$ 的 x 所组成的集合 $\sqrt{H}$, 那么 $\sqrt{H}$ 是一个理想.
   (b) 证明：如果 H 是多项式环 $\mathbf{C}[x,y,z]$ 中的一个理想, V 是相应的轨迹, 那么 $J(V)$ 包含 $\sqrt{H}$. (*提示*：希尔伯特零点定理断言 $J(V)=\sqrt{H}$.)
8. 描述下面两种情形下, 由 $x^2+y^2=0$ 所确定的轨迹.
   (a) 在 $\mathbf{R}^3$ 中.　　(b) 在 $\mathbf{C}^3$ 中.

*13.6 线性代数中的理想

在非交换环中, 我们可以考虑"单边"理想. 环 A 中一个子集合 L, 如果 x 和 y 在 L 中, a 在 A 中, 则 $x-y$ 与 ax 也在 L 中, 那么称 L 是 A 的*左理想*. 右理想可以类似地定义. 与这些概念相对照, 我们原来意义下的理想称为*双边理想*. 例如, 在所有 2×2 矩阵组成的环 M_2 中, 第一列都是零的所有矩阵构成左理想, 但不能构成双边理想.

这些概念可以有效地应用到含有单位元素 1 的线性代数 A 中; 正像 13.1 节中看到的, 任意这样的线性代数是一个环. 在这种情况下, 任意左理想 L 或右理想关于数乘运算也是封闭的. 例如, 如果 ξ 是 L 中任意元素, c 是任意标量, 那么 L 包含 $c\xi$, 这因为 $c\xi=(c\cdot 1)\xi$ 是 L 中元素与 A 中某元素 $c\cdot 1$ 的乘积. 如果把 A 看作它的标量域 F 上的线性空间, 那么 A 的任意左 (或右) 理想是一个子空间.

如果一个线性代数没有真 (双边) 理想, 则称它为*单代数*. 于是一个单代数没有真同态像.

定理 10 域上所有 $n\times n$ 矩阵组成的代数是单代数.

证明 这个代数 M_n 以 n^2 个矩阵 $\boldsymbol{E}_{ij}$ 作为它的基, $\boldsymbol{E}_{ij}$ 是在 (i,j) 位置的元素为 1, 其余位置都是零. M_n 中的一个真理想 B 将至少包含一个非零矩阵 $\boldsymbol{A}=\sum\limits_{i,j}\alpha_{ij}\boldsymbol{E}_{ij}$(其中系数 $\alpha_{rs}\neq 0$). 那么每个矩阵

$$(\alpha_{rs})^{-1}\boldsymbol{E}_{kr}\boldsymbol{A}\boldsymbol{E}_{sk}=(\alpha_{rs})^{-1}\sum_{i,j}\boldsymbol{E}_{kr}\boldsymbol{E}_{ij}\boldsymbol{E}_{sk}\alpha_{ij}=\boldsymbol{E}_{kk} \tag{26}$$

在 B 中. 因此单位矩阵 $I=\sum\limits_{k}\boldsymbol{E}_{kk}$ 在 B 中, 所以 B 一定是整个代数, 于是 B 是假理想. 证毕

范德波恩 (Wedderburn)(1908) 证明了有名的定理 10 的逆定理. 这个逆定理断言, 特别地, 复数域 $\mathbf{C}$ 上的每个单代数与 $\mathbf{C}$ 上所有 $n\times n$ 矩阵组成的代数同构. 为了讨论一般情形, 我们需要可除代数的概念. 可除代数是指这个线性代数是一个可除环. 根据代数基本定理我们可以证明, 复数域 $\mathbf{C}$ 上唯一的可除代数是 $\mathbf{C}$ 本身. 著名的弗罗比尼乌斯定理断言, 实数域 $\mathbf{R}$ 上的可除代数只有 $\mathbf{R}$, $\mathbf{C}$ 和四元数代数 (8.11 节).

在任意可除环 D 上, 我们可以构造一个任意 n 阶全矩阵代数 $M_n(D)$, 如下所述. 可以按照普通法则

$$\begin{aligned}(a_{ij})+(b_{ij})&=(a_{ij}+b_{ij}),\\ c(a_{ij})&=(ca_{ij}),\\ (a_{ij})(b_{ij})&=\left(\sum_{k=1}^{n}a_{ik}b_{kj}\right),\end{aligned} \tag{27}$$

把系数在可除代数D 中的两个 $n\times n$ 矩阵相加和相乘. 范德波恩的结果是, 如果 F 是任意域, 域 F 上最一般的单代数 A 可如下得到. 取域 F 上任意可除代数 D 和任意正整数 n, 那么 A 是由系数在 D 中的所有 $n\times n$ 矩阵组成.

习 题

1. 证明: 每个可除代数是单代数.
2. 在可除代数中求出所有右理想.
3. 讨论一个环中左理想构成的代数, 即描述它们的和、交与主左理想.
4. 证明: 域 F 上线性代数的每个商环本身是一个线性代数.
5. (a) 证明: 如果 S 是向量空间 F^n 的子空间, 以 S 中的向量为行的所有矩阵组成的集合是 $M_n(F)$ 的左理想.

*(b) 证明: $M_n(F)$ 的每个左理想 C 是 (a) 中所描述的一个理想. (提示: 证明 C 的矩阵的每一行是 C 中除了第一行外剩下的其他行都为零的矩阵的第一行. 利用 7.6 节和 7.7 节的方法.)

*6. 把定理 10 推广到任意可除环 D 上的全矩阵代数 $M_n(D)$.

13.7 环的特征

任意环 R 可以看作加法群 (阿贝耳群). 由任意 $a \in R$ 生成的循环子群是由 a 的 m 次幂组成, 其中 m 取整数. 用加法的记号, 我们把 a 的 m 次 "幂" 写成 $m \times a$. 于是, 如果 m 是正整数, 则

$$m \times a = a + a + \cdots + a \qquad (m\text{ 个被加项}). \tag{28}$$

如果 $m = 0$, 则 $0 \times a = 0$, 而当 $m = -n$ 是负数时, 则有

$$(-n) \times a = n \times (-a) = (-a) + (-a) + \cdots + (-a)\ (n\text{ 个被加项}). \tag{29}$$

我们称 $m \times a$ 是 a 的 m 次自然倍数, 它对任意 $m \in \mathbf{Z}$ 和任意 $a \in R$ 都有定义.

整环 D 中的元素的这些自然倍数具有任意交换环中幂的所有性质, 这些性质已经在 6.6 节中证明过, 那里是按照乘法记号来叙述的. 因此有

$$(m \times a) + (n \times a) = (m + n) \times a, \quad m \times (n \times a) = (mn) \times a, \tag{30}$$

$$m \times (a + b) = m \times a + m \times b, \quad m \times (-a) = (-m) \times a. \tag{31}$$

还有由分配律推出的一些性质. 其中一个一般分配律 (见 1.5 节) 是

$$(a + a + \cdots + a)b = ab + ab + \cdots + ab \qquad (m\text{ 个被加项}).$$

按照自然倍数表示这就是

$$(m \times a)b = m \times (ab) = a(m \times b). \tag{32}$$

当 $m = 0$ 时, 这个公式仍成立, 当 m 是负数时公式也成立, 这因为取 $m = -n$, 定义 (29) 给出

$$(-n) \times ab = n \times (-ab) = [n \times (-a)]b = [(-n) \times a]b.$$

另一个一般分配律是法则

$$(a + \cdots + a)(b + \cdots + b) = ab + \cdots + ab.$$

它也可以改写成

$$(m \times a)(n \times b) = (mn) \times (ab). \tag{33}$$

对一切整数 m 和 n, 不管正的、负的或零, 这个公式都成立.

令 $a = 1$ 是 R 的单位元素 (乘法单位元素), (32) 表明 $m \times b$ 正是 $(m \times 1)b$, 它是 b 与 1 的 m 次自然倍数的乘积. 此外, 在 (30) 中令 $a = 1$, 我们看出, 从 $\mathbf{Z}$ 到 R 的映射 $m \longmapsto m \times 1$ 保持和. 最后, 在 (33) 中令 $a = b = 1$, 我们得到

$$(m \times 1)(n \times 1) = (mn) \times (1 \times 1) = (mn) \times 1. \tag{33'}$$

这个映射保持积. 这就证明了

定理 11 对任意环 R, 映射 $m \longmapsto m \times 1$ 是从 $\mathbf{Z}$ 到 R 的一个同态.

推论 1 任意环 R 中 1 的自然倍数组成的集合是一个同构于 $\mathbf{Z}$ 或 $\mathbf{Z}_m$($m > 1$ 为某整数) 的子环.

定义 环 R 的特征是指 R 的单位元素 1 的不同自然倍数 $m \times 1$ 的个数 m.

推论 2 在整环 D 的加法群中, 所有非零元素具有相同的阶 —— 即 D 的特征.

证明 对于所有非零元素 $b \in D, m \times b = 0$ 当且仅当 $(m \times 1)b = 0$, 根据消去律, 这等价于 $m \times 1 = 0$. 证毕

整数环 $\mathbf{Z}$ 具有特征[①]∞, 而整环 $\mathbf{Z}_p$ 具有特征 p. ∞ 和 p 是唯一可能的特征.

定理 12 一个整环的特征或者是 ∞ 或者是素数 p.

用反证法证明定理. 我们假定某个整环 D 有有限特征, 这个特征是一个复合数 $m = rs$, 那么根据 (33′), 环 D 的单位元素 1 满足

$$0 = m \times 1 = (rs) \times 1 = (r \times 1) \cdot (s \times 1).$$

根据消去律, 或者 $r \times 1 = 0$, 或者 $s \times 1 = 0$. 因此 D 的特征一定是其中一个因子或者 r, 或者 s, 而不是我们所假定的 m.

推论 在任意整环中, 由单位元素生成的加法群是一个与 $\mathbf{Z}$ 或 $\mathbf{Z}_p$ 同构的子整环.

1.5 节的二项公式 (9) 说明了自然倍数的意义. 在任意交换环 R 中, 表达式

$$(a+b)^2 = a^2 + ab + ba + b^2 = a^2 + 2 \times (ab) + b^2$$

有一个中项, 确切地说它是自然倍数 $2 \times (ab)$. 更一般地, 用归纳法可以证明 1.5 节中的二项公式 (9), 它包含的二项系数是自然倍数, 于是我们可以写成

$$(a+b)^n = a^n + \binom{n}{1} \times (a^{n-1}b) + \binom{n}{2} \times (a^{n-2}b^2) + \cdots + \binom{n}{n} \times b^n, \tag{34}$$

① 大多数作者用 "特征 0" 来代替 "特征 ∞".

其中系数 $\binom{n}{i}$ 是按公式

$$\binom{n}{i}=\frac{n!}{(n-i)!i!},\quad i=0,1,\cdots,n \tag{35}$$

给出的自然数. 这里 $n!=n(n-1)\cdots 3\times 2\times 1$, 并且 $0!=1$.

定理 13 *在素特征 p 的任意交换环 R 中, 对应 $a\longmapsto a^p$ 是一个同态.*

证明 根据 (6), 我们需要证明 $1^p=1$, 并且对所有 $a,b\in R$ 有 $(ab)^p=a^pb^p,(a\pm b)^p=a^p\pm b^p$. 前两个方程在任意交换环中都成立. 为证明第三个方程, 在公式 (34) 和 (35) 中, 令 $n=p$. 因为 p 是素数, 所以它不能被 $i!$ 或 $(p-i)!(0<i<p)$ 的任意因子整除. 因此 (34) 中适合 $0<i<p$ 的所有二项系数都是 p 的倍数. 但是环 R 具有特征 p, 因此 (34) 中所有含 $\binom{p}{i}(0<i<p)$ 的项被去掉. 由此推出等式

$$(a\pm b)^p=a^p\pm b^p. \tag{36}$$

定理证毕.

推论 *在特征为 p 的有限域 F 中, 对应 $a\longmapsto a^p$ 是一个自同构.*

证明 因为在 F 中由 $a^p=0$ 可推出 $a=0$, 所以同态 $a\longmapsto a^p$ 的核是零, 而且这个同态是一一的. 因为 F 是有限的, 所以这就意味着 $a\longmapsto a^p$ 还是映上的. 因此是一个自同构.

习　题

1. 证明: 对于正整数 m, 自然倍数 $m\times a$ 可以通过"递推公式" $1\times a=a,(m+1)\times a=m\times a+a$ 来定义.
2. 用归纳法证明关于正自然倍数的法则 (30) 和 (32).
3. 作为定理 13 的一个推论来证明费马定理 (1.9 节定理 18).
4. 关于有序整环的特征, 你能说些什么?
5. (a) 证明: 在特征为 p 的任意整环 D 中, $\alpha:a\longmapsto a^p$ 是一一的 (单一同态).
   (b) 证明: 如果 $D=\mathbf{Z}_p(x)$, 那么 α 的像是 D 的真子整环.
   (c) 证明: 有限域必有一个真自同构, 除非它是一个素域 $\mathbf{Z}_p$.

13.8 域的特征

因为域被定义为除法 (除零外) 是可能的整环, 所以关于特征的讨论立刻可应用到域上. 如果域 F 的特征为 p, 那么根据定理 12, 由域 F 的单位元素生成的加法子群是一个子域, 并且它与由模 p 整数构成的有限域同构. 如果域 F 的特征为 ∞,

那么根据定理 12, 由单位元素 1 生成的子群是由所有倍数$m\times 1$ 组成, 所以由 c 生成的子域是由所有商$\dfrac{m\times 1}{n\times 1}$(其中 $n\neq 0$) 组成. 这个子域是所有倍数 $m\times 1$ 的子整环的商域. 因此根据 2.2 节的定理 7, 它与有理数域同构, 有理数域是整数 $m\longleftrightarrow m\times 1$ 构成的整环的商域. 事实上, 映射 $\dfrac{m\times 1}{n\times 1}\longleftrightarrow\dfrac{m}{n}$ 是由 1 生成的子域与有理数域之间的同构. 这就证明了下面的结果 (参看 2.6 节定理 18 的推论 2):

定理 14 *在特征为 ∞ 的域中, 由单位元素生成的子域与有理数域 $\mathbf{Q}$ 同构.*

同构$\dfrac{m\times 1}{n\times 1}\longleftrightarrow\dfrac{m}{n}$ 保持域 F 中的所有四种运算. 于是在处理单个域 F 时, 可以把每个商 $\dfrac{m\times 1}{n\times 1}$ 与其相应的 $\dfrac{m}{n}$ 等同起来. 在这个约定之下, 可以说每个特征为 ∞ 的域包含着所有有理数 $\dfrac{m}{n}(n\neq 0)$. 按照类似的约定, 可以说每个特征为 p 的域包含域 Z_p. 在这种意义下, 每个域都是极小域(即所谓素域)$\mathbf{Q}$ 和 $\mathbf{Z}_p$ 中的一个域的扩张. 因此按照对已知域的不同扩张方式对域进行系统的分类是很自然的. 这些将在下一章里讨论.

习　题

1. 设 F_4 是恰有四个元素的任意域,
   (a) 证明: F_4 的特征为 2.
   (b) 证明: F_4 的不在素子域 $\mathbf{Z}_2$ 中的两个元素都满足 $x^2=x+1$.
   (c) 用这个事实证明: F_4 与 13.3 节中习题 8 所述的域 $\mathbf{Z}[\omega]/(2)$ 同构.
2. 求出习题 1 的域 F_4 的全部自同构.
3. 证明: 关于二次方程的解的一般公式可以应用到特征为 2 的任意域上.
4. 在什么样的域上, 求解三次方程的一般公式 (5.5 节) 仍然成立.

第14章 代数数域

14.1 代数扩张与超越扩张

剩下的两章讨论一般域 F 上的多项式方程 $p(x)=0$ 的解及其性质. 我们将证明任意这样的方程在 F 的适当的扩张中是可解的, 这个扩张是指包含 F 作为子域的一个域 K. 例如 $p(x)=0$ 在多项式环 $F[x]$ 对于由 p 的倍式组成的主理想的商域 $F[x]/(p)$ 中总有一个根.

在描述这种扩张的一般性质之后, 我们将特别地研究有理数域 $\mathbf{Q}$ 按这种方式扩张而得到的所有"代数数"构成的域. 通过证明在某一个二次扩张 $\mathbf{Q}[x]/(x^2-r)=q(\sqrt{r})(r\in\mathbf{Z})$ 中"整数"的唯一因子分解定理的问题, 简短地介绍一下代数数论. 例如, 高斯整数 $m+n\sqrt{-1}(r=-1$ 的情形) 可以唯一地分解成高斯素数.

域 F 最简单的一类扩张 K 是由单个元素 $c\in K$ 的有理表达式 $\dfrac{p(c)}{q(c)}=\dfrac{\Sigma a_kc^k}{\Sigma b_lc^l}$(系数 $a_k,b_l\in F$) 组成. 例如, 复数 $a+bi$ 是由实数和单个复数 i 生成, 而未定元 x 的所有有理形式 (具有有理系数) 组成的域 $\mathbf{Q}(x)$ 是由域 $\mathbf{Q}$ 和元素 x 生成的. 一个域可以按照几种不同方法生成. 例如, 域 $\mathbf{Q}(\sqrt{2})$ 由方程 $x^2-2=0$ 的根 $\sqrt{2}$ 生成, 它是由含有有理系数 a 和 b 的所有实数 $a+b\sqrt{2}$ 组成 (见 2.1 节的例子). 另一个不同的方程 $x^2+4x+2=0$ 的一个根是 $-2+\sqrt{2}$, 它生成同一个域 $\mathbf{Q}(\sqrt{2})$, 因为这个域中的任意数可以按照这个新的生成元表示为

$$a+b\sqrt{2}=(a+2b)+b(-2+\sqrt{2}).$$

普通的配方方法应用到这个方程上得到 $x^2+4x+2=(x+2)^2-2$, 所以 $y=x+2$ 满足新的方程 $y^2-2=0$, 其根生成同一个域, 于是运用变量替换来化简方程对应着在相应的域中选取新的生成元.

我们来一般地描述由域 F 的任意扩张 K 中的已知元素生成的子域. 设 K 是给定的域, F 是 K 的子域, c 是 K 的一个元素. 考虑 K 中那些由形为

$$f(c)=a_0+a_1c+a_2c^2+\cdots+a_nc^n \quad (\text{每个 } a_i\in F) \tag{1}$$

的多项式给出的元素. K 中包含 F 和 c 的任意子整环一定包含所有这样的元素 $f(c)$. 反过来, 所有这样的多项式组成的集合在加法、减法和乘法运算之下是封闭的. 因此这些表达式 (1) 组成 K 中由 F 和 c 生成的子整环. 这个子整环一般都用带方括号的 $F[c]$ 来表示.

如果 $f(c)$ 和 $g(c)\neq 0$ 是像 (1) 那样的多项式表达式, 那么它们的商 $\dfrac{f(c)}{g(c)}$ 是 K 的元素, 称为具有系数属于 F 的关于 c 的*有理表达式*. 所有这样的商组成的集合是一个子域; 它是由 F 和 c 生成的域, 一般都用带*圆括号*的 $F(c)$ 来表示.

如果域 K 是在它的子域 F 上由单个元素 c 生成的, 则称 K 是 F 的*单扩张*, 所以 $K=F(c)$. 2.1 节所讨论的域 $\mathbf{Q}(\sqrt{2}),\mathbf{Q}(\sqrt[3]{5})$ 和 $\mathbf{Q}(\omega)$ 都是单扩张的例子. 可以证明, 不管怎样, F 的任意扩张可以通过单扩张的一个有限或 (良序) 超限序列而得到.

在有理数域上, 一些复数 (如 i), $\sqrt{2},\sqrt[3]{5},\sqrt{-3}$ 都满足有理系数多项式方程. 还有另外一些数, 像 π 和 e=2.718 28⋯, 可以证明它们不满足有理系数多项式方程 (平凡情形除外). 后面这些数称为"超越数". 这种重要的二分法可以应用到任意域的元素上.

定义 *设 K 是任意域, F 是 K 的任意子域. 域 K 的元素 c, 如果满足一个多项式方程*

$$a_0+a_1c+a_2c^2+\cdots+a_nc^n=0 \quad (a_i\in F,\text{且不全为零}), \tag{2}$$

则称 c 在 F 上是代数的. K 的元素 c 如果在 F 上不是代数的, 则称 c 在 F 上是超越的.

单扩张 $K=F(c)$ 是 F 上的代数扩张还是超越扩张, 取决于生成元 c 在 F 上是代数的还是超越的. 单超越扩张的结构是特别容易描述的.

定理 1 *如果 c 在 F 上是超越的, 那么由 F 和 c 生成的子域 $F(c)$ 与系数在 F 中的关于未定元 x 的所有有理形式组成的域 $F(x)$ 同构. 同构可以选为 $a\longmapsto a$(每个 $a\in F$), $c\longmapsto x$.*

证明 扩张 $F(c)$ 显然包含 F 和系数在 F 中的所有有理表达式 $\dfrac{f(c)}{g(c)}$. 如果 $F(c)$ 中的两个多项式表达式 $f_1(c)$ 和 $f_2(c)$ 相等, 那么它们的系数一定逐项相等, 因为如果不然, 差 $f_1(c)-f_2(c)$ 将产生一个关于 c 的系数不全为零的多项式方程, 这与 c 在 F 上是超越的假设相矛盾. 因此对应 $f(c)\longmapsto f(x)$ 是整环 $F[c]$ 和未定元 x 的多项式形式整环 $F[x]$ 之间的双射. 根据多项式运算法则, 这个对应是一个同构. 根据 2.2 节定理 6, 它可以扩张成 $F(c)$ 和 $F(x)$ 之间的同构 $\dfrac{f(c)}{g(c)}\longleftrightarrow\dfrac{f(x)}{g(x)}$.

习 题

1. 辨别下列各复数在有理数域 $\mathbf{Q}$ 上是代数的还是超越的, 并给出证明: $\sqrt{7},\sqrt[3]{5},\pi^2,\mathrm{e}+3$(这里 $\mathrm{e}=2.718\ 28\cdots$), i+3, $\mathrm{e}^{2\pi\mathrm{i}},\sqrt{2}+\mathrm{i}$.
2. 证明: 如果 x 在 F 上是代数的, 那么 x^2 和 $x+3$ 也是代数的, 反之亦然.
3. $\mathbf{Q}(\sqrt{5})$ 中什么样的数生成整个域 $\mathbf{Q}(\sqrt{5})$?

4. (a) 设 d 是非完全平方整数, 描述域 $\mathbf{Q}(\sqrt{d})$.
   (b) 求出 $\mathbf{Q}(\sqrt{d})$ 中生成整个域的那些元素.
   (c) 把每个这样的元素表示为系数在 $\mathbf{Q}$ 中的二次方程的根.

14.2 域上的代数元素

下面我们研究域 F 的单代数扩张的性质, 该扩张是由 F 和在 F 上单个的代数元素 u 生成的. 根据定义, 这个元素必满足 F 上次数至少是 1 的多项式方程. 同一个元素 u 可以满足很多不同的方程, 例如, $\sqrt{2}$ 是 $x^2-2=0$ 的根, 也是 $x^3-2x=0, x^4-4=0$ 等方程的根. 但是它恰好是一个首一不可约多项式方程的根 (见后面的习题 6).

定理 2 *如果域 F 的扩张 K 的一个元素 u 在 F 上是代数的, 那么 u 是多项式整环 $F[x]$ 中唯一的一个首一不可约多项式 $p(x)$ 的零点. 如果 h 是 $F[x]$ 中另一个多项式, 那么 $h(u)=0$ 当且仅当在整环 $F[x]$ 中, h 是 p 的倍式, 也就是当且仅当 h 在 $F[x]$ 的主理想 (p) 中.*

证明 满足 $h(u)=0$ 的多项式 $h \in F[x]$ 组成 $F[x]$ 中的一个理想, 这个理想恰好是 "赋值映射" $p \longmapsto p(u)$ 所定义的同态 $\phi_u: F[x] \to K$ 的核, 这里映射 $p \longmapsto p(u)$ 在 $u \in K$ 处赋给每个多项式 p 一个值. 像 $F[x]$ 的所有理想一样, 这个理想是主理想 (3.8 节定理 11), 所以它由它的任意一个次数最低的元素的所有倍式组成. 这些次数最低的元素中恰有一个是首一多项式, 称它为 p. 这个 p 是不可约的, 如果不然, 它可以分解为 $p=fg$, 这里 f 和 g 是次数更小的多项式, 由此推出 $f(u)g(u)=p(u)=0$, 所以或者 $f(u)=0$, 或者 $g(u)=0$, 这与选取 p 为适合 $p(u)=0$ 的次数最低的多项式相矛盾. 证毕

定义 *在域 F 上代数元素 u 的极小多项式是满足 $p(u)=0$ 的 (唯一的) 首一不可约多项式 $p \in F[x]$. u 在 F 上的次数 $n=[u:F]$ 是这个多项式的次数.*

推论 *如果元素 u 在 F 上的次数为 n. 那么我们有, $a_0+a_1u+\cdots+a_{n-1}u^{n-1}=0(a_i \in F)$ 当且仅当 $a_0=a_1=\cdots=a_{n-1}=0$.*

现在我们有可能描述 K 的由 F 和上述的代数元素 u 所生成的子域. 这个子域 $F(u)$ 显然包含由所有可表为系数在 F 中的多项式 $f(u)$ 的元素组成的子整环 $F[u]$. 此外还将证明, 映射 $f(x) \longmapsto f(u)$ 是商环 $F[x]/(p)$ 与 $F(u)$ 之间域的同构 $\phi': F[x]/(p) \to F(u)$.

这一节余下的部分将讨论这个结论. 由多项式的加法和乘法公式, 显然有 ϕ' 是由 $F[x]$ 到子整环 $F[u]$ 的一个满同态. 但是实际上, 整环 $F[u]$ 是一个子域. 事实上, 让我们求 $F[u]$ 中任意元素 $f(u) \neq 0$ 的逆. 不等式 $f(u) \neq 0$ 意味着 u 不是 $f(x)$ 的根, 因此根据定理 2, $f(x)$ 不是不可约多项式 $p(x)$ 的倍式, 所以 $f(x)$ 与 $p(x)$ 互素.

因此我们可以写

$$1 = t(x)f(x) + s(x)p(x), \tag{3}$$

这里 $t(x)$ 和 $s(x)$ 是 $F[x]$ 中适当的多项式. $F[u]$ 中相应的方程是 $1 = t(u)f(u)$. 这就表明, $F[u]$ 的非零元素 $f(u)$ 有一个逆元素 $t(u), t(u)$ 也是 u 的多项式①, 于是就证明了 $F[u]$ 是 K 的子域.

反之, 因为 K 的每个包含 F 和 u 的子域显然包含 $F[u]$ 中的每个多项式 $f(u)$, 所以我们看出 $F[u]$ 是 K 的由 F 和 u 生成的子域. 我们证明了

定理 3 *设 K 是任意域, u 是 K 的一个元素, 它在 K 的子域 F 上是代数的; 设 $p(x)$ 是以 u 为其根的 F 上首一不可约多项式. 那么从多项式整环 $F[x]$ 到 $F(u)$ 的映射 $\phi' : f(x) \longmapsto f(u)$ 是以 $(p(x))$ 为核的满同态.*

把这个定理同 13.3 节定理 5 的推论 2 结合起来, 我们有一个直接理论.

定理 4 *在定理 3 中, $F(u)$ 与商环 $F[x]/(p)$ 同构, 这里 p 是 u 所满足的域 F 上首一不可约多项式.*

商环 $F[x]/(p)$ 可以描述得非常简单. 每个多项式 $f(x) \in F[x]$ 在模 (p) 之下与它用 $p(x)$ 除所得的余式 $r(x) = f(x) - a(x)p(x)$ 同余, 这个余式是次数小于 n 的唯一的多项式

$$r(x) = r_0 + r_1 x + \cdots + r_{n-1}x^{n-1}. \tag{4}$$

把两个这样的多项式相加或相减, 恰好是对它们的相应的系数相加或相减. 为把它们相乘, 先按照 3.1 节的 (3′) 计算出多项式乘积, 然后再计算用 $p(x)$ 除时所得的余式.

例如, 有理数域 $F = \mathbf{Q}$ 通过 $u = \sqrt{2}$ 扩张成 $\mathbf{Q}(\sqrt{2})$, 在这个特殊情形中, 我们有 $p(x) = x^2 - 2$. 因此 $\mathbf{Q}(u)$ 的任意元素可以写成 $a + b\sqrt{2}$, 其中 a 和 b 为有理数, 并且

$$\begin{aligned}(a + b\sqrt{2})(c + d\sqrt{2}) &= ac + (ad + bc)\sqrt{2} + bd(\sqrt{2})^2 \\ &= (ac + 2bd) + (ad + bc)\sqrt{2}.\end{aligned}$$

公式 (4) 表明商环 $F[x]/(p)$ 是 F 上一个 n 维向量空间, 它是有限维向量空间 $F[x]$ 对由 $p(x)$ 的倍式组成的子空间的商空间. 还要注意, 乘法是双线性的 (对每个因子是线性的). 因此代数扩张 $F[x]/(p)$ 在 13.1 节的意义下也可以看作 F 上的交换线性代数.

---

① 例如, 在 $\mathbf{Q}(\sqrt{3})$ 中, $1+\sqrt{3}$ 有乘法逆, 通过分母有理化可以求得这个逆是

$$\frac{1}{1+\sqrt{3}} = \frac{1-\sqrt{3}}{(1+\sqrt{3})(1-\sqrt{3})} = -\frac{1}{2} + \frac{1}{2}\sqrt{3}.$$

习 题

1. 求出 $\sqrt{3}$ 所满足的五个不同多项式方程, 并证明它们都是 $\sqrt{3}$ 所满足的首一不可约多项式 (域 $\mathbf{Q}$ 上) 的倍式.
2. 在由不可约方程 $u^3-6u^2+9u+3=0$ 的根 u 生成的单扩张 $\mathbf{Q}(u)$ 中, 按照元素 1, u, u^2 把下列各元素表示成 (4) 中的形式:
$$u^4, u^5, 3u^5-u^4+2, \frac{1}{u+1}, \frac{1}{u^2-6u+8}.$$
3. 在由 $x^5+2x+2=0$ 的根 u 生成的单扩张 $\mathbf{Q}(u)$ 中, 把下列各元素表示成(4) 中的形式: $(u^3+2)(u^3+3u), u^4(u^4+3u^2+7u+5), \frac{1}{u}, \frac{u+2}{u^2+3}$.
4. 把复数域表示为由所有实系数多项式组成的整环 $\mathbf{R}[x]$ 得到的商环.
5. 把域 $\mathbf{Q}(\sqrt{2})$ 表示为由有理系数多项式组成的整环 $\mathbf{Q}[x]$ 得到的商环.
6. 根据有关的定义直接证明: 如果 u 在 F 上是代数的, 那么以 u 为根的次数最低的首一多项式在 F 上是不可约的.
7. 根据有关的定义证明: 如果 u 是域 K 的任意元素, F 是 K 的任意子域, 那么以 u 为根的, 系数在 F 中的所有多项式 $g(x)$ 的集合是 $F[x]$ 的一个理想.

14.3 根的添加

迄今, 我们假定已经给出域 F 的扩张 K, 并描述了 K 的由 F 和已知元素 $u\in K$ 生成的子域, 这里 u 是由 F 上的极小多项式 (即首一不可约多项式) p 给出的, 它满足 $p(u)=0$. 另一方面, 我们恰恰可以从 F 和不可约多项式 p 出发, 构造一个包含 $p(x)=0$ 的根的较大的域. 这种构造方法是把第 5 章用过的由实数域 $\mathbf{R}$ 添加上方程 $x^2+1=0$ 的一个虚根来构造复数域 $\mathbf{C}$ 的方法加以一般化. 定理 3 和定理 4 指出在一般情形下怎样得到同样的结果.

定理 5 *如果 F 是域, p 是 F 上不可约多项式, 那么存在域 $K\cong F[x]/(p)$, 它是由 $p(x)$ 的根 u 生成的 F 的单代数扩张.*

证明 因为 $p(x)$ 是不可约的, 所以主理想 (p) 是 $F[x]$ 中的极大理想. 因此根据 13.3 节定理 6, 商环 $F[x]/(p)$ 是一个域, 它包含 F 和剩余类 $x+(p)$(该剩余类包含 x), 并满足在 $F[x]/(p)$ 中有 $p(x)=0$.

这个单扩张除同构外是唯一的.

定理 6 *如果域 $F(u)$ 和 $F(v)$ 是同一域 F 的两个单代数扩张, 它们分别由 F 上同一个不可约多项式 p 的根 u 和 v 生成, 那么 $F(u)$ 与 $F(v)$ 同构. 特别恰好存在一个 $F(u)$ 到 $F(v)$ 的同构, 在这个同构之下, u 对应于 v, F 的每个元素与自身对应.*

证明 由定理 4 提供的同构
$$F(u)\xleftarrow{\phi_u}F[x]/(p)\xrightarrow{\phi_v}F(v)$$

取它们的合成 $\phi_u^{-1}\phi_v$.

定理 5 可以用来构造各种有限域. 例如, 从模 3 整数的域 $\mathbf{Z}_3$ 出发, 对于多项式 x^2-x-1, $0,1,2$ 三个元素没有一个是它的零点, 因此它在 $\mathbf{Z}_3[x]$ 中是不可约的. 所以商环 $\mathbf{Z}_3[x]/(x^2-x-1)$ 是一个域 K, 它是由它的子域 $\mathbf{Z}_3$ 和 x 的陪集 (称为 u) 生成的. 而且因为 $[u:F]=2$, 所以这个域 K 的每个元素可以唯一地写成 $a+bu$, 其中 $a,b\in F$, 因此 K 恰好有 9 个元素.

这个域还可以不用商环概念直接来构造. 它刚好由 9 个形为 $a+bu$ 的元素组成. 它们之中两个元素之和由法则

$$(a+bu)+(c+du)=(a+c)+(b+d)u$$

给出. 为计算两个这种类型的元素之积, 我们可先按自然方式乘出来, 然后再根据已给出的方程 $u^2=u+1$ 来化简. 其结果是

$$(a+bu)(c+du)=ac+(ad+bc)u+bdu^2=(ac+bd)+(ad+bc+bd)u.$$

我们可以详细验证, 这 9 个元素 $a+bu(a,b\in\mathbf{Z}_3)$ 在上述两种运算之下满足域的所有公设. 特别是, 非零元素的逆由下表给出:

1	2	u	$2u$	$1+u$	$1+2u$	$2+u$	$2+2u$
1	2	$2+u$	$1+2u$	$2+2u$	$2u$	u	$1+u$

根据上述构造, 这个域显然是由剩余类域 $\mathbf{Z}_3$ 添加 u 生成的域 $\mathbf{Z}_3(u)$. 它是有限域中最简单的例子之一 (见 15.3 节).

上述添加方式可以用到任意基域 F 上. 如果 F 是实数域 $\mathbf{R}$, $p(x)$ 是 $\mathbf{R}$ 上不可约多项式 x^2+1, 那么这个构造得到域 $\mathbf{R}(u)$, 它是由满足 $u^2=-1$ 的数 u 生成. 这个数 u 的性质很像 $\mathrm{i}=\sqrt{-1}$, 并且域 $\mathbf{R}(u)$ 实际上与复数域 $\mathbf{C}$ 同构, 这同我们在第 5 章中用过的从实数域得到复数域的构造方法稍微有些不同.

如果 F 是模 p 整数的域 $\mathbf{Z}_p$, $p(x)$ 是 F 上某一不可约多项式, 则上面的构造方法将产生一个由元素 $a_0+a_1u+\cdots+a_{n-1}u^{n-1}$ 组成的域. 因为每个系数 a_i 只有 p(有限) 种选择, 因此这样构造出的域是具有 p^n 个元素的有限域, 这里 n 是多项式 $p(x)$ 的次数.

用同样的方法, 我们还可以构造代数函数域. 例如, 设 $F=\mathbf{C}(z)$ 是所有有理复函数组成的域, 假设我们要求把满足 $t^2=(z^2-1)(z^2-4)$ 的函数 $t(z)$ 添加到 F 上. 我们可以把多项式 $p(t)=f(z,t)=t^2-(z^2-1)(z^2-4)$ 看作系数在 $\mathbf{C}(z)$ 中的 t 的二次不可约多项式. 那么商环 $K=F[t]/(p(t))$ 是一个包含所有有理函数和代数函数 t 的域. 我们可以把 $t(z)$ 作为 K 的一个元素来研究, 而不必对它 (它是双值的) 构造黎曼面. 域 K 称为*椭圆函数域*, 因为它是由椭圆积分

$$\int\sqrt{(z^2-1)(z^2-4)}\mathrm{d}z$$

的被积函数生成的.

如果把定理 6 应用到像 x^3-5 这样的普通多项式 (它在有理数域 $\mathbf{Q}$ 上不可约) 上, 可以得到由正的 $\sqrt[3]{5}$ 生成的 $\mathbf{Q}$ 的扩张 $\mathbf{Q}(\sqrt[3]{5})$, 也可得到扩张 $\mathbf{Q}(\omega\sqrt[3]{5})$, 这里 $\omega=\dfrac{-1+\mathrm{i}\sqrt{3}}{2}$ 是复三次单位根. 可以证明这两个域 $\mathbf{Q}(\sqrt[3]{5})$ 和 $\mathbf{Q}(\omega\sqrt[3]{5})$ 在代数上没有什么区别, 因为它们是同构的.

粗略地说, 这个同构意味着一个不可约多项式 $p(x)$ 的任意两个根具有相同的性质, 根 u 的所有代数性质都可以从它所满足的不可约方程推导出来. 有很多这样的同构例子. 例如, 复数域 $\mathbf{C}=\mathbf{R}(\mathrm{i})$ 是在实数域 $\mathbf{R}$ 上添加方程 $x^2+1=0$ 的两个根 $\pm\mathrm{i}$ 中任意一个而生成的, 因此根据定理 6, 存在一个把 i 映射到 $-\mathrm{i}$ 的 $\mathbf{C}$ 的自同构. 这个自同构刚好是一个数和它的通常共轭复数之间的对应 $a+b\mathrm{i}\longleftrightarrow a-b\mathrm{i}$.

习　题

1. 列出下列各域的非恒等对应的自同构: $\mathbf{Q}(\sqrt{2}),\mathbf{Q}(\sqrt{-3}),\mathbf{Q}(\mathrm{i})$.
2. 分别列出一个由复数组成的非实域与实域 $\mathbf{Q}(\sqrt[3]{5})$ 和 $\mathbf{Q}(\sqrt[4]{2})$ 同构.
3. 证明: x^3+x-1 在模 5 整数域 $\mathbf{Z}_5$ 上是不可约的. 如果把这个多项式的根添加到 $\mathbf{Z}_5$ 上, 那么所得到的域有多少个元素?
4. (a)[①] 求出在模 2 整数域 $\mathbf{Z}_2$ 上不可约的 2 次和 3 次多项式.
   (b) 对四元素域构成加法表和乘法表.
5. (a) 证明: 正文中构造的九元素域的特征是 3.
   (b) 对这个域明显地列出同构 $a\longleftrightarrow a^3$.
6. (a) 求出域 $\mathbf{Z}_3$ 上所有二次不可约多项式.
   (b) 证明: 任意两个九元素域同构. (提示: 首先证明这样的有限域中每个元素在 $\mathbf{Z}_3$ 上是二次的.)
7. 证明: t 的多项式 $t^2-(x^2-1)(x^2-4)$ 在域 $\mathbf{C}(x)$ 上是不可约的. (提示: 用 3.9 节的结果.)
8. 证明: 正文中的椭圆函数域 $\mathbf{C}(x,y)$ 可以通过在 $\mathbf{C}(x)$ 上添加方程 $t^2=\dfrac{x^2-4}{x^2-1}$ 的根而得到.
9. 如果 $g(t)$ 是可约多项式, 那么在商环 $F[t]/(g(t))$ 中哪些元素确有逆元素?
10. 用 13.3 节定理 6 给出关于 $F[t]/(p(t))$ 是域的另一个证明.

14.4　次数与有限扩张

在一个 n 次元素 u 生成的单扩张 $F(u)$ 中, 每个元素 w 按公式 (4) 有唯一的表达式为

① 原书此题有误, 现按第 3 版译出. —— 译者注

$$w = a_0 + a_1u + \cdots + a_{n-1}u^{n-1}, \tag{5}$$

其中系数在 F 中. 这唯一的表达式同一个向量按照基向量 $1, u, \cdots, u^{n-1}$ 的表达式极为相似. 这就暗示我们运用向量空间的概念. 的确, 域 F 的任意扩张可以看作域 F 上的向量空间：只要不管域 K 的元素的乘法, 而把 K 的两个元素相加和 K 的元素同 F 的元素的"数乘"两种运算当作向量空间的运算, 这些加法和数乘运算满足向量空间的所有公设. 如果这个向量空间 K 是有限维的, 那么称域 K 是 F 的有限扩张, 并且把这个向量空间的维数 n 称为扩张次数 $n = [K : F]$.

例如, 复数域 $\mathbf{C} = \mathbf{R}(\mathrm{i})$ 是实子域 $\mathbf{R}$(像在 5.2 节中那样) 上的二维向量空间; 由有理数域 $\mathbf{Q}$ 和 5 的三次根生成的域 $\mathbf{Q}(\sqrt[3]{5})$ 是有理子域 $\mathbf{Q}$ 上的三维向量空间, 等等. 一般地, 关于单代数扩张的定理 4 可以按照维数重述如下.

定理 7　*域 F 上代数元素 u 的次数, 等于把扩张 $F(u)$ 看作 F 上向量空间时 $F(u)$ 的维数. 这个向量空间有一组基 $1, u, \cdots, u^{n-1}$.*

在 14.5 节中, 我们将要说明如何用向量空间的方法来分析由域 F 添加几个不同代数元素而得到的扩张. 但是在讨论这样的"多重扩张"之前, 我们首先来看一下这种向量空间的方法怎样能使我们比较 F 的同一个单代数扩张 $F(u)$ 中的不同元素所满足的不可约方程.

关于向量空间的一个基本事实是维数的不变性 (向量空间的任意两组基元素的个数相同). 这个事实可以应用到域的有限扩张这种特殊情形, 如下所述.

推论　*如果域 F 上的两个代数元素 u 和 v 生成同一个扩张 $F(u) = F(v)$, 那么 u 和 v 在 F 上的次数相同.*

一个单代数扩张是有限扩张, 反之, 每个有限扩张是由代数元素组成的.

定理 8　*F 的有限扩张 K 的每个元素 w 在 F 上是代数的, 并且满足一个次数至多是 n 的 F 上不可约方程, 这里 $n = [K : F]$ 是给定的扩张的次数.*

证明　给定元素 w 的 $n+1$ 个幂 $1, w, w^2, \cdots, w^n$ 是 n 维向量空间 K 的元素, 因此在 F 上一定线性相关 (7.4 节定理 5 的推论 2). 所以必有线性关系 $b_0 + b_1w + \cdots + b_nw^n = 0$, 其中系数不全为零. 可以把它解释为多项式, 于是这个关系就意味着 w 在 F 上是代数的.

推论　*单代数扩张 $F(u)$ 的每个元素在 F 上是代数的.*

这个重要的结论使我们确信, 超越元素绝不能出现在一个单代数扩张中.

在讨论一个特殊的单代数扩张 $F(u)$ 时, 要系统地应用 u 所满足的不可约多项式 $p(x)$, 根据定理 2, 这个扩张中的元素 $g(u)$ 是零当且仅当多项式 $g(x)$ 可被 $p(x)$ 整除. 例如, 假定 $\mathbf{Q}(u)$ 是有理数域 $\mathbf{Q}$ 上的三次扩张, 它是由 $x^3 - 2x + 2$ 的一个根 u 生成. 根据爱森斯坦不可约准则 (3.10 节), 这个多项式是不可约的. 这个扩张 $\mathbf{Q}(u)$ 中的元素 $w = u^2 - u$ 一定满足某个次数至多是 3 的多项式方程. 为求出这个方程, 像

在定理 4 中那样, 按照 1, u 和 u^2 把幂 $w^2 = u^4 - 2u^3 + u^2, w^3 = u^6 - 3u^5 + 3u^4 - u^3$ 线性地表示出来. 反复运用已知的方程 $u^3 = 2u - 2$, 这是可以做到的. 由此得到

$$w = u^2 - u, \quad w^2 = 3u^2 - 6u + 4, \quad w^3 = 16u^2 - 28u + 18.$$

$1, w, w^2$ 和 w^3 之间一定满足一个线性关系, 为得到这个关系, 我们可以由前两个线性方程解出 u 和 u^2 为

$$u = -\frac{w^2}{3} + w + \frac{4}{3}, \qquad u^2 = -\frac{w^2}{3} + 2w + \frac{4}{3}. \tag{6}$$

把这些代入 w^3 的表达式中就得到所需要的方程

$$w^3 - 4w^2 - 4w - 2 = 0.$$

根据爱森斯坦定理, 这个方程在 $\mathbf{Q}$ 上是不可约的, 根据方程 (6) 我们也可以说 u 在 $\mathbf{Q}(w)$ 中, 所以 $\mathbf{Q}(u) = \mathbf{Q}(w)$, 于是 u 和 w 生成同一个扩张, 根据定理 7 的推论可知, 它们在 $\mathbf{Q}$ 上的次数都是 3. 这就意味着 w 所满足的任意三次方程一定是不可约的.

习　题

1. 下列每个数都在 $\mathbf{Q}$ 的一个单代数扩张中, 因此在 $\mathbf{Q}$ 上是代数的. 对每种情形, 求出该数所满足的首一不可约方程.
   (a) $2 + \sqrt{3}$,　(b) $\sqrt[4]{5} + \sqrt{5}$,　(c) $\sqrt[3]{2} + \sqrt[3]{4}$,
   (d) $u^2 - 1$, 这里 u 满足 $u^3 = 2u + 2$,　(e) $u^2 + u$, 这里 u 满足 $u^3 = -3u^2 + 3$.
2. 证明: 实数域 $\mathbf{R}$ 的每个有限扩张或者是 $\mathbf{R}$ 本身, 或者同构于复数域 $\mathbf{C}$.
3. 证明: 复数域没有真有限扩张.
4. (a) 证明: 如果 K 是有理数域 $\mathbf{Q}$ 的二次扩张, 那么 $K = \mathbf{Q}(\sqrt{d})$, 其中 d 是一个整数, 非完全平方且无平方因子.
   (b) 如果域 $\mathbf{Q}$ 用特征为 ∞ 的域 F 代替, 那么上述结果还正确吗? 如果用任意特征的域代替, 情况如何?
5. 未定元 x 的有理形式构成的域 $F(x)$ 是 F 的有限扩张吗? 为什么?
6. 证明: 特征为 p 的有限域中元素的个数是 p 的幂.
7. (a) 证明: 在模 p 整数域 $\mathbf{Z}_p$ 上恰有 $\dfrac{p^2 - p}{2}$ 个二次首一不可约多项式.
   (b) 证明: 对每个 p, 都存在一个含有 p^2 个元素特征为 p 的域.

*8. 证明: 在模 p 整数域 $\mathbf{Z}_p$ 上恰有 $\dfrac{p^3 - p}{3}$ 个三次首一不可约多项式.

*9. 设 F 是包含在整环 D 中的一个域, 证明:
   (a) D 是 F 上的向量空间.
   (b) 如果把 D 看作 F 上的向量空间, 它是有限维的, 那么 D 是一个域.

14.5 多重代数扩张

域 F 的有限扩张可以通过一系列的单扩张来构造. 如果 F 的特征是 ∞, 那么我们可以证明, 任意这样的多重扩张可以表示成一个单扩张, 也就是说, 它是在 F 上添加一个适当选择的单个元素而生成的. 我们将略去此证明, 而直接来讨论多重扩张的性质. 一般地, 如果 K 是 F 的包含元素 $c_1, c_2, \cdots, c_r$ 的任意扩张, 那么记号 $F(c_1, c_2, \cdots, c_r)$ 表示由 $c_1, \cdots, c_r$ 和 F 的元素生成的 K 的子域 (这个子域是由系数在 F 中关于 $c_1, \cdots, c_r$ 的有理形式所表示的所有元素组成的). 另一方面, 这样的多重扩张可以通过反复进行单扩张而得到. 例如, $F(c_1, c_2)$ 是单扩张 $L = F(c_1)$ 的单扩张 $L(c_2)$.

在求解方程时可以产生多重代数扩张, 在这里引进适当的辅助方程常常是有用的. 例如, 方程 $x^4 - 2x^2 + 9 = 0$ 可以写成

$$x^4 - 2x^2 + 9 = (x^4 - 6x^2 + 9) + 4x^2 = (x^2 - 3)^2 + 4x^2 = 0.$$

所以这个方程变为 $\left(\dfrac{x^2-3}{2x}\right)^2 = -1$. 这个公式表明, 包含上述给定方程的根 u 的任意域, 也包含方程 $y^2 = -1$ 的根 $\mathrm{i} = \dfrac{u^2-3}{2u}$. 如果我们把辅助量 i 添加到有理数域 $\mathbf{Q}$ 上, 那么原来的方程在 $\mathbf{Q}(\mathrm{i})$ 上就变为可约的, 因为

$$x^4 - 2x^2 + 9 = (x^2 - 3 + 2x\mathrm{i})(x^2 - 3 - 2x\mathrm{i}).$$

根据普通的公式, 因式 $x^2 - 3 - 2\mathrm{i}x$ 有一个根 $u = \mathrm{i} + \sqrt{2}$. 于是原来方程在域 $K = \mathbf{Q}(\sqrt{2}, \mathrm{i})$ 中就有一个根 u. 这个域 K 可以在 $\mathbf{Q}$ 上先添加 $\sqrt{2}$, 后添加 i 而得到. 中间域 $\mathbf{Q}(\sqrt{2})$ 是由实数组成的, 因此不可能包含 i. 所以 i 所满足的二次方程 $y^2 + 1 = 0$ 在实数 $\mathbf{Q}(\sqrt{2})$ 上一定仍然是不可约的, 所以扩张 $\mathbf{Q}(\sqrt{2}, \mathrm{i})$ 在 $\mathbf{Q}(\sqrt{2})$ 上的次数是 2, 它的两个基元素是 1 和 i. 而域 $\mathbf{Q}(\sqrt{2})$ 在 $\mathbf{Q}$ 上有一组基 1, $\sqrt{2}$. 所以在整个域 $\mathbf{Q}(\sqrt{2}, \mathrm{i})$ 中的任意元素 w 可以表示成

$$w = (a + b\sqrt{2}) + (c + d\sqrt{2})\mathrm{i} = a + b\sqrt{2} + c\mathrm{i} + d\sqrt{2}\mathrm{i}, \tag{7}$$

其中 a, b, c, d 是有理数. 于是 $1, \sqrt{2}, \mathrm{i}, \sqrt{2}\mathrm{i}$ 这四个元素构成 $\mathbf{Q}$ 上整个扩张 $K = \mathbf{Q}(\sqrt{2}, \mathrm{i})$ 的一组基. 这种计算基的方法可以一般地叙述如下:

定理 9 如果元素 $u_1, \cdots, u_n$ 构成 F 的有限扩张 K 的一组基, 而 $w_1, \cdots, w_m$ 组成 K 的扩张 L 的一组基, 那么 mn 个乘积 $u_i w_j (i = 1, \cdots, n; j = 1, \cdots, m)$ 构成 F 的扩张 L 的一组基.

证明 L 中任意元素 y 可以表示成给定基的线性组合 $y=\sum_j r_jw_j$, 这里系数 $r_j\in K$. 每个系数 r_j 又可以表示成 K 的这组基元素的某线性组合 $r_j=\sum_i a_{ij}u_i$, 这里每个系数 $a_{ij}\in F$. 代入这些值, 得到

$$y=\sum_j\sum_i a_{ij}u_iw_j,$$

这表现为已假定的元素 u_iw_j 的一个线性组合, 这里系数在 F 中. 用同样类型的逐次论证方法可以证明, 这 mn 个元素在 F 上是线性无关的, 因此它们构成 L 的一组基. 证毕

由定理 9 可以得出很多推论. 首先, 我们可以把与所用的特殊基无关的结果叙述如下:

推论 1 *如果 K 是 F 的有限扩张, L 是 K 的有限扩张, 那么 L 是 F 的有限扩张, 它的次数是*

$$[L:F]=[L:K][K:F]\quad(L\supset K\supset F).\tag{8}$$

推论 2 *如果 K 是 F 的次数为 $n=[K:F]$ 的有限扩张, 那么 K 的每个元素 u 在 F 上的次数是 n 的因子.*

证明 元素 u 生成单扩张 $F(u)$, 因此根据 (8) 式有 $n=[K:F(u)][F(u):F]$, 这里第二个因子是我们所考虑的元素 u 的次数.

推论 3 *有限扩张 $K\supset F$ 的元素 u 生成整个扩张当且仅当 $[K:F]=[u:F]$.*

证明 如果 u 在 F 上满足一个次数为 $[K:F]$ 的不可约方程, 那么 u 在 F 上生成 n 次子域 $F(u)$. 根据 (8) 式, 这个子域一定包含整个 K.

推论 4 *如果 $K=F(y_1,y_2,\cdots,y_r)$ 是由 r 个量 $y_1,\cdots,y_r$ 生成的域, 其中逐个 y_i 在由前 $i-1$ 个量生成的域 $F(y_1,\cdots,y_{i-1})$ 上是代数的, 那么 K 是 F 的有限扩张, K 中的每个元素在 F 上是代数的.*

证明 每个次数 $[F(y_1,\cdots,y_{i-1},y_i):F(y_1,\cdots,y_{i-1})]$ 是有限的, 因此根据推论 1, 整个次数 $[K:F]$ 是有限的, 根据定理 8, K 中每个元素在 F 上是代数的.

推论 5 *如果 $p(x)$ 是域 F 上一个三次不可约多项式, K 是 F 的 2^m 次扩张, 那么 $p(x)$ 在 K 上是不可约的.*

这个推论特别意味着, 三次不可约方程绝不能通过逐次求平方根的方法来解, 这是因为, 把一个平方根添加到域 F 上, 或者全然没有给出扩张, 或者给出二次扩张, 所以由任意多个平方根得到的扩张 $K=F(\sqrt{a},\sqrt{b},\sqrt{c},\cdots)$ 的次数是 2 的某个幂 2^m. 根据推论 5, 这个扩张绝不包含给定三次不可约方程的根.

为证明推论, 假定 $p(x)$ 在 2^m 次的域 K 上是可约的. 那么三次多项式 $p(x)$ 一定至少有一个线性因子 $x-u$, 于是 K 包含 $p(x)$ 的根 u. 但是根据推论 2, 这种在

F 上的次数为 3 的元素 u 不可能包含在 F 上的次数为 2^m 的域 K 中. 这就证明了 $p(x)$ 在 K 上是不可约的.

这个推论是下述定理的代数基础：只用直尺和圆规不可能解一般的倍立方或三等分任意角这样的经典问题. 任意这样的作图问题可以化为解析的形式. 这个问题的对象是由一些点和直线组成. 对于某一组坐标轴, 这些点的坐标 (和这些直线的方程中系数之比) 是一个实数集合, 它生成某个由实数组成的域 F. 在用直尺和圆规作图时, 每一步都提供某些新的点和直线. 可以证明①, 相应的新的数域或者是 F 本身或者是 F 的二次扩张, 因此重复上述作图便产生点和直线的集合, 它对应于 F 上的 2^m 次的域 K.

现在考虑倍立方问题. 问题的对象是由三个坐标轴, 每个轴上的一个单位线段, 以这些线段为边的一个立方体组成. 这个问题是作另外一个具有两倍体积的立方体. 这个新立方体的边长将满足方程 $x^3-2=0$. 根据爱森斯坦定理, 这个方程在有理数域 $\mathbf{Q}$ 上是不可约的 (这个域 $\mathbf{Q}$ 与问题的对象有联系). 根据推论 5, 在对应于直尺和圆规作图的域 K 上, 多项式 x^3-2 还是不可约的. 因此通过这种方法不可能构造出 (比如说沿 x 轴) 一个线段, 使它是倍立方体的边.

三分角问题可按类似的方法处理, 问题的实质在于, 写出一个用整个角的余弦来表示三分之一角的余弦的三角方程. 对于大多数角来说, 这又给出一个三次不可约方程.

习　题

1. 在定理 9 中, 详细证明：mn 个元素 u_iw_j 在 F 上线性无关.
2. 证明：正文中所处理的方程 x^4-2x^2+9 在 $\mathbf{Q}$ 上是不可约的. (提示：用域 $\mathbf{Q}(\sqrt{2},\mathrm{i})$ 的次数证明.)
3. 证明：如果 $p(x)$ 是 F 上 q 次不可约多项式, K 是 F 的有限扩张, 其次数与 q 互素, 那么 $p(x)$ 在 K 上是不可约的.
4. 确定有理数域 $\mathbf{Q}$ 上的下列各多重扩张的次数, 并说明理由.
   (a) $\mathbf{Q}(\sqrt{3},\mathrm{i})$,　(b) $\mathbf{Q}(\sqrt[3]{5},\sqrt{-2})$,　(c) $\mathbf{Q}(\sqrt{18},\sqrt[4]{2})$,　(d) $\mathbf{Q}(\sqrt{8},3+\sqrt{50})$,
   (e) $\mathbf{Q}(\sqrt[3]{2},u)$, 这里 u 满足 $u^4+6u+2=0$,　(f) $\mathbf{Q}(\sqrt{3},\sqrt{-5},\sqrt{7})$,　(g) $\mathbf{Q}(\sqrt{3},\sqrt{2})$.
5. 对习题 4 中的每个域给出 $\mathbf{Q}$ 上的一组基.
6. 确定下列多项式在指定的域上是否是不可约的, 并给出理由.
   (a) x^2+3, 在 $\mathbf{Q}(\sqrt{7})$ 上;　(b) x^2+1, 在 $\mathbf{Q}(\sqrt{-2})$ 上;
   (c) x^3+8x-2, 在 $\mathbf{Q}(\sqrt{2})$ 上;　(d) x^5+3x^3-9x-6, 在 $\mathbf{Q}(\sqrt{7},\sqrt{5},1+\mathrm{i})$ 上.
7. 在下列每种情形中, 确定给出的数 u 是否生成给出的有理数域 $\mathbf{Q}$ 的扩张. 对每种情形证明你的答案是正确的.

① 本质上, 这依赖于下述事实：圆 (圆规) 的方程是二次的, 直线 (直尺) 的方程是线性的.

(a) $u=\sqrt[3]{7}$, 在 $\mathbf{Q}(\sqrt[3]{7})$ 中; (b) $u=\sqrt{2}+\sqrt{5}$, 在 $\mathbf{Q}(\sqrt{2},\sqrt{5})$ 中;

(c) $u=2+\sqrt[3]{9}$, 在 $\mathbf{Q}(\sqrt[3]{3})$ 中; (d) $u=\dfrac{\sqrt{2}-1}{1+\sqrt{2}}$, 在 $\mathbf{Q}(\sqrt{2})$ 中;

(e) $u=v^2+v+1$, 在 $\mathbf{Q}(v)$ 中, 这里 v 满足 $v^3+5v-5=0$.

8. $c=\pi^6+5\pi^3+2\pi-14$ 在有理数域 $\mathbf{Q}$ 上是超越的还是代数的? 为什么?
9. 证明: 如果 K 是 F 的素数次扩张, 那么不在 F 中的 K 中任意元素在 F 上生成整个 K.
10. (a) 求用 $\cos 3\theta$ 给出 $\cos\theta$ 的三次方程.

    (b) 证明: 当 $3\theta=60°$ 时, 这个方程在 $\mathbf{Q}$ 上是不可约的 (这意味着 60° 角不能用直尺和圆规三等分).

14.6 代 数 数

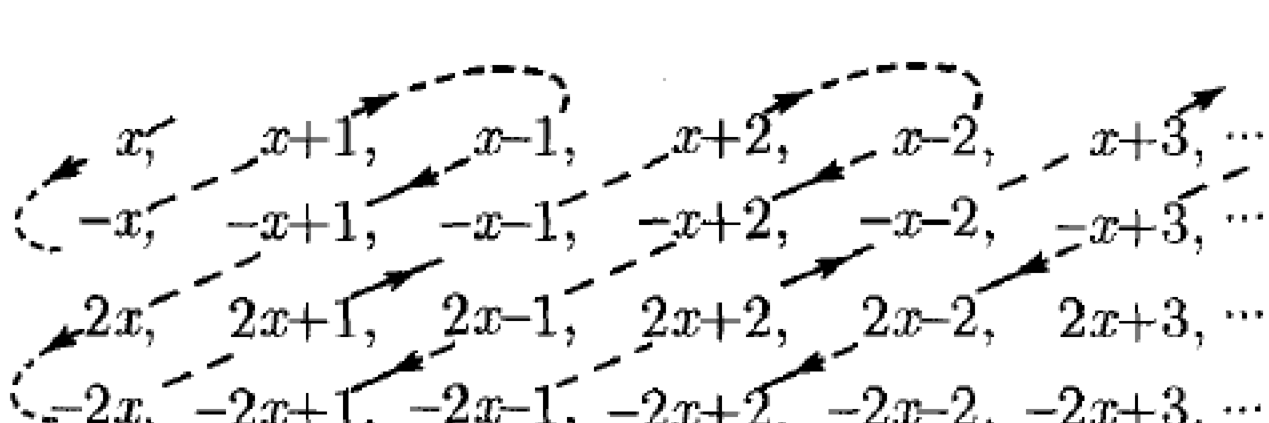
代数数 u 是满足含有不全为零有理系数的多项式方程的复数. 即

$$a_0+a_1u+a_2u^2+\cdots+a_nu^n=0,\quad (a_i\in\mathbf{Q}, a_i \text{ 不全为零}). \tag{9}$$

换句话说, 代数数是在有理数域 $\mathbf{Q}$ 上代数的任意复数. 在讨论域的扩张时, 我们反复用到代数数的例子, 例如 i, $\sqrt{-2}$, $\sqrt[5]{3}$ 或 ω.

定理 10 所有代数数组成的集合是可数的.

验证这个命题需要我们描述一下对所有代数数进行计数或排列的方法. 首先, 我们把它们所满足的方程都列出来. 我们注意, 代数数所满足的方程 (9) 可以用它的有理系数的公分母去乘, 于是得到一个含有不全为零的整系数的方程, 可以假定这个方程的首项系数是正的. 我们知道这些多项式的所有可能的整系数是可数的, 例如列为 $0,+1,-1,+2,-2,+3,-3,\cdots$. 全体整系数线性多项式可以排成一个阵列, 如



那么我们可以按照箭头所示的顺序陆续取出上面阵列的对角线元素, 把所有线性多项式排成一个单列, 其结果是

$$x,-x,x+1,x-1,-x+1,2x,-2x,2x+1,-x-1,\cdots.$$

然后我们再求出二次多项式的长方形阵列, 这只要把各种二次项 mx^2 添加到上述单列中的每个元素上. 由此阵列我们又可得到一个包含所有二次多项式的单列, 对

高次多项式也如此去做. 当对每个次数的多项式都按此法做完之后, 结果得到由这些单列组成的阵列, 其中第 n 行是包含所有 n 次多项式的单列. 再取这个阵列的对角线元素, 把它展开, 我们就得到包含所有多项式的一个单列. 在此单列中, 每个多项式都用它的根来代替, 并去掉那些重复的根. 这个结果就是包含所有整系数多项式的根的单列, 这也就是说, 所有代数数是可数的.

定理 10 的一个推论是：全体实代数数是可数的. 但是康托对角线法证明了 (12.3 节定理 5) 所有实数组成的集合是不可数的. 因此实数集合大于所有实代数数组成的集合. 这个论证对实超越数的存在性给出一个间接证明. 我们把这个结果叙述如下：

推论　*不是每个实数都是代数数.*

最初有很多数学家不相信康托的论证, 因为这个论证没有给出任何特殊的实超越数. 但是现在, 他的论证已被广泛接受, 并有可能对这个推论给出更明显的证明.

定理 11　*所有代数数组成的集合是一个域.*

证明　我们只须证明, 任意两个代数数 $u, v \neq 0$ 的和、积、差、商仍然是代数数. 但是所有这些组合都包含在由 u 和 v 生成的复数域的子域 $\mathbf{Q}(u,v)$ 中. 因为 u 在 $\mathbf{Q}$ 上是代数的, 所以 $\mathbf{Q}(u)$ 是 $\mathbf{Q}$ 的有限扩张; 因为 v 在 $\mathbf{Q}(u)$ 上是代数的, 所以 $\mathbf{Q}(u,v)$ 是 $\mathbf{Q}(u)$ 的有限扩张. 因此根据定理 9, $\mathbf{Q}(u,v)$ 是 $\mathbf{Q}$ 的有限扩张, 于是它的每个元素是代数数 (定理 8).　　证毕

如果系数在 F 中的每个多项式方程的根在 F 中, 那么称域 F 是代数完全的[①]. 在这样的域 F 上每个多项式 $f(x)$ 有一个根 c, 因此 $f(x)$ 有线性因子 $x-c$. 所以 F 上仅有的一类不可约多项式是线性的, 因而代数完全域 F 上每个多项式都可以写成线性因子的乘积 (像 5.3 节公式 (11) 中所表示那样). 进一步, 除了 F 本身之外, 不可能有 F 的单代数扩张. 于是我们得出结论：域 F 是代数完全的当且仅当 F 没有真单代数扩张. 代数基本定理 (5.3 节定理 5) 断言, 复数域是代数完全的.

定理 12　*所有代数数组成的域 A 是代数完全的.*

证明　取多项式方程 $x^n + u_{n-1}x^{n-1} + \cdots + u_0 = 0$, 它的系数 u_i 都是 A 中的代数数. 这些系数生成一个扩张 $K = \mathbf{Q}(u_0, u_1, \cdots, u_{n-1})$, 根据定理 9 的推论 4, 它是有理数域 $\mathbf{Q}$ 的有限扩张. 给定的这个方程的任意复根 r 在域 K 上是代数的, 所以 $K(r)$ 是 K 的有限扩张, 因而也是 $\mathbf{Q}$ 的有限扩张. 根据定理 8, 这个扩张中的元素 r 则在 $\mathbf{Q}$ 上是代数的. 这就意味着根 r 是一个代数数, 所以也就在 A 中, 因此 A 是代数完全的.　　证毕

我们现在把域 $\mathbf{Q}$ 嵌入到由所有代数数组成的代数完全域 A 中, 把实数域 $\mathbf{R}$ 嵌入到由所有复数组成的代数完全域 $\mathbf{C}$ 中. 这些结果是下述一般定理的特殊情形, 这

① 有些情况下用 “代数闭 (algebraically closed)” 来代替 “代数完全 (algebraically complete)”. 但考虑到与拓扑学类比, 用 “代数完全的” 这个术语比较好.

个定理指出, 任意域 F 不管怎样都有一个扩张 A, 这个 A 是代数完全的, 并且 A 中每个元素在 F 上是代数的 (参看 15.1 节的附录).

代数数论已发展得很完整, 它主要研究代数数组成的域 K, 它是有理数域 $\mathbf{Q}$ 的有限扩张. 这样的域称为代数数域. 我们后面考虑这个域的算术性质.

习 题

1. 通过求出下列每个代数数所满足的有理系数方程来说明定理 11:
   (a) $\sqrt{2}+\sqrt{-3}$, (b) $\sqrt{-1}+\sqrt[3]{5}$, (c) $(\sqrt{7})(\sqrt[3]{2})$,
   (d) $\dfrac{\sqrt{7}}{1+\sqrt{2}}$, (e) $u\sqrt{-2}$, 这里 u 满足 $u^3+7u-14=0$.
2. (a) 证明: 如果 u 和 v 分别是 ($\mathbf{Q}$ 上) m 次和 n 次代数数, 则 $u+v$ 的次数不超过 mn.
   (b) $\dfrac{u}{v}$ 的次数怎样?
   (c) 证明: 如果 t 是超越数, u 是代数数, 那么 $t+u$ 和 tu 都是超越数, 对后一种情形, 我们假定 $u\neq 0$.
3. 通过求出下列每个方程的根所满足的有理系数方程来说明定理 12.
   (a) $x^2+3x+\sqrt{2}=0$,
   (b) $x^2+\sqrt{3}x-\sqrt{-1}=0$,
   (c) $x^3-\sqrt{3}x+1+\sqrt[3]{2}=0$,
   (d) $x^2+u+2=0$, 这里 u 是方程 $u^3+5u^2-10u+5=0$ 的根.
4. 像定理 10 的证明中所指出的那样, 列出所有二次多项式所组成的单列的前 16 项.
5. 不用定理 10 证明: 固定次数的所有代数数组成的集合是可数的.
6. 证明: 可数域的任意有限扩张是可数的.
7. 证明: 在域 F 的任意可数子域 S 上是代数的所有元素组成的集合 A 是可数的.
8. (a) 证明: 存在一个实数, 它在 $\mathbf{Q}(\pi)$ 上是超越的.
   (b) 用习题 7 和 3.4 节的定义证明: 存在可数多个代数无关的实数.
9. 指出在定理 10 的证明中隐含地用到下面的超限算术公式:
   (a) 存在 $\mathbf{d}^{n+1}=\mathbf{d}$ 个 n 次多项式.
   (b) 存在 $\mathbf{d}+\mathbf{d}+\cdots+\mathbf{d}+\cdots$(到 $\mathbf{d}$ 项)$=\mathbf{d}^2$ 个多项式 (所有次数).

*10. (a) 设 u 是任意固定的实数, 通过把 x^i-u^i 因式分解来证明: 存在常数 $N(j)$, 使得不管是否有 $|x-u|<1$, 都有 $|x^j-u^j|\leqslant N|x-u|$.
   (b) 设 $f(x)$ 是实系数多项式, u 是任意实数, 证明: 存在一个与 f 和 u 有关的常数 M, 使得不管是否有 $|x-u|<1$, 都有 $|f(x)-f(u)|\leqslant M|x-u|$.

*11. 设实代数数 u 满足 r 次整系数多项式方程 $f(x)=0$. 证明: 如果 m 和 n 是适合 $\left|\dfrac{m}{n}-u\right|<\dfrac{1}{Mn^r}$(其中 M 是习题 10 中的常数), 那么 $f\left(\dfrac{m}{n}\right)=0$. (提示: 用习题 10, 有 $\left|f\left(\dfrac{m}{n}\right)\right|<\dfrac{1}{n^r}$, 而 $f\left(\dfrac{m}{n}\right)$ 是分母为 n^r 的有理数.)

*12. 证明: 如果 u 是实数, 对于 u 可以求出不同有理数的无穷序列 $\dfrac{m_k}{n_k}$ 满足$\left|u-\dfrac{m_k}{n_k}\right|<$

$\frac{1}{kn_k^k}$(对所有的 k), 那么 u 是超越数. $\Big($提示：如果 u 的次数是 r, 那么由习题 11 得到, 对所有充分大的 k, 有 $f\left(\frac{m_k}{n_k}\right) = 0.\Big)$

*13. 满足习题 12 的假设条件的数称为刘维尔 (Liouville) 数 (超越数).

(a) 证明：$\sum_{k=1}^{\infty} 10^{-k!} = 0.110\,001\cdots$ 是刘维尔数.

(b) 再列出两个别的刘维尔数.

14.7 高斯整数

高斯整数是分量 a, b 都为整数的复数 $\alpha = a + b\mathrm{i}$. 任意这样的高斯整数满足整系数的首一多项式方程 $\alpha^2 - 2a\alpha + (a^2 + b^2) = 0$, 因此它是代数数. 两个高斯整数的和、差、积仍然是高斯整数, 因此全体高斯整数构成整环 $\mathbf{Z}[\mathrm{i}]$. 在这个整环中可以考虑可除性和分解成素 (不可约) 因子的问题.

对任意复数 σ, 引进 "范数" (是整数, 或者不是) 是方便的. 设 $\sigma = r + s\mathrm{i}$, 则范数$N(\sigma)$ 是 σ 与它的共轭复数 $\sigma^* = r - s\mathrm{i}$ 的乘积：

$$N(\sigma) = \sigma\sigma^* = (r + s\mathrm{i})(r - s\mathrm{i}) = r^2 + s^2. \tag{10}$$

这个范数永远是非负的, 并且是 σ 的绝对值的平方. 对任意两个复数 σ 和 τ, 我们有

$$N(\sigma\tau) = N(\sigma)N(\tau). \tag{11}$$

这个等式意味着, 对应 $\sigma \longmapsto N(\sigma)$ 保持乘积, 换句话说, 它是由非零的数 σ 组成的乘法群到实数的乘法群上的同态映射. 特别是, 高斯整数的范数是 (有理) 整数[①].

现在回忆一下包含可除性的一般概念 (3.6 节). $\mathbf{Z}[\mathrm{i}]$ 的单位是这样的高斯整数：$\alpha \neq 0$, 它的倒数 α^{-1} 也是一个高斯整数. 那么有 $\alpha\alpha^{-1} = 1$, 所以 $N(\alpha\alpha^{-1}) = N(\alpha)N(\alpha^{-1}) = 1$, 因而单位 α 的范数一定是 $N(\alpha) = 1$. 由 (10) 式可以看出, $\mathbf{Z}[\mathrm{i}]$ 的单位只能是 ± 1 和 $\pm\mathrm{i}$. 如果 $\mathbf{Z}[\mathrm{i}]$ 中两个高斯整数可彼此整除, 则称这两个整数在 $\mathbf{Z}[\mathrm{i}]$ 中相伴. 因此在 $\mathbf{Z}[\mathrm{i}]$ 中 α 的相伴只能是 $\pm\alpha$ 和 $\pm\mathrm{i}\alpha$.

有理素数 5 在 $\mathbf{Z}[\mathrm{i}]$ 中有四种不同的分解

$$\begin{aligned} 5 &= (1 + 2\mathrm{i})(1 - 2\mathrm{i}) = (2\mathrm{i} - 1)(-2\mathrm{i} - 1) \\ &= (2 + \mathrm{i})(2 - \mathrm{i}) = (\mathrm{i} - 2)(-\mathrm{i} - 2). \end{aligned} \tag{12}$$

这些分解本质上是一样的, 例如, $2 + \mathrm{i} = \mathrm{i}(1 - 2\mathrm{i})$ 和 $2 - \mathrm{i} = -\mathrm{i}(1 + 2\mathrm{i})$, 其他各种情形中相应的因子也都是相伴. (12) 中的每个因子都是素的[②](不可约的). 例如, 如果

① 有理整数即普通整数, 定义见下节. —— 译者注

② 即不可分解的, 其定义见后, 它与有理素数即普通素数的定义不同. —— 译者注

2+i 可以分解为 $2+\mathrm{i}=\alpha\beta$, 那么 $N(2+\mathrm{i})=5=N(\alpha)N(\beta)$, 所以 $N(\alpha)$(或者 $N(\beta)$) 等于 1, 因此 α(或者 β) 是单位. (12) 式的各个因子本质上只给出 5 的一个因子分解, 因为在任意分解 $5=\gamma\delta$ 中, $N(5)=25=N(\gamma)N(\delta)$, 所以每个不是单位的因子一定有范数 5. 通过试验我们发现, 范数为 5 的所有高斯整数就是 (12) 式中写出的那些.

另一方面, 有理素数 3 在 $\mathbf{Z}[\mathrm{i}]$ 中是素的. 假定 $3=\alpha\beta$, 则 $N(\alpha)N(\beta)=9$, 于是 $N(\alpha)|9$. 如果 $N(\alpha)=1$, 则 α 是单位; 如果 $N(\alpha)=N(a+b\mathrm{i})=3$, 则 $a^2+b^2=3$, 不可能有整数 a 和 b 使这个等式成立. 因此在高斯整数环中 3 没有真因子 α.

高斯整数的唯一因子分解定理可以通过构造除法算式来证明, 这个算式类似于对普通整数和多项式用过的除法算式.

定理 13 *对于给定的高斯整数 α 和 $\beta\neq 0$, 存在高斯整数 γ 和 ρ 适合*

$$\alpha=\beta\gamma+\rho,\quad N(\rho)<N(\beta). \tag{13}$$

证明 我们从商 $\dfrac{\alpha}{\beta}=r+s\mathrm{i}$ 出发, 选取 r' 和 s' 是与有理数 r 和 s 最接近的整数. 那么

$$\frac{\alpha}{\beta}=(r'+s'\mathrm{i})+[(r-r')+(s-s')\mathrm{i}]=\gamma+\sigma,\quad \gamma=r'+s'\mathrm{i},$$

这里 $|r-r'|\leqslant\dfrac{1}{2},|s-s'|\leqslant\dfrac{1}{2}$, 所以

$$N(\sigma)=(r-r')^2+(s-s')\leqslant\frac{1}{4}+\frac{1}{4}<1.$$

上述方程现在可写成 $\alpha=\beta\gamma+\beta\sigma$, 其中 $\alpha,\beta\gamma$ 都是高斯整数, 因此 $\beta\sigma$ 也是高斯整数, 而且这里 $N(\beta\sigma)=N(\beta)N(\sigma)<N(\beta)$. 证毕

引理 1 *两个高斯整数 α_1 和 α_2 有最大公因子 δ, 它是一个可表示成 $\delta=\beta_1\alpha_1+\beta_2\alpha_2$ 形式的高斯整数, 其中 β_1 和 β_2 都是高斯整数.*

证明 通过辗转相除, 我们可以构造欧几里得算法, 它很像有理整数情形 (1.7 节). (13) 式的逐次余数 ρ 的范数越来越小, 因此这种算法最终会结束. 最后的非零余数就是我们所要求的最大公因子. 证毕

另一个证明是从环 $\mathbf{Z}[\mathrm{i}]$ 中的 α_1 和 α_2 生成的理想 (α_1,α_2) 出发. 在这个理想的全体元素中间选择一个范数最小的元素 δ, 并像 (13) 式那样, 写 $\alpha_1=\delta\gamma_1+\rho_1,\alpha_2=\delta\gamma_2+\rho_2$. 这些余数 ρ_i 都在理想中, 并且它们的范数小于 δ 的范数, 因此一定是零. 所以 $\alpha_1=\delta\gamma_1,\alpha_2=\delta\gamma_2$, 于是 δ 是公因子. 因为 δ 在这个理想中, 所以它有形式 $\delta=\beta_1\alpha_1+\beta_2\alpha_2$, 因此它是 α_1 和 α_2 的每个公因子的倍数, 所以 δ 就是我们所要求的最大公因子.

高斯整数分解的其余部分可同有理整数的情形 (1.7 节和 1.8 节) 及多项式的情形 (3.5 节和 3.8 节) 完全一样地进行处理, 因此我们这里只叙述重要的步骤. 一个高斯整数 π, 如果它不是 0 也不是单位, 而且它在 $\mathbf{Z}[\mathrm{i}]$ 中的因子只能是单位和 π 的相伴, 则称 π 是素的. 我们可以证明

引理 2 如果 π 是素的, 那么由 $\pi|\alpha\beta$ 可推出 $\pi|\alpha$ 或者 $\pi|\beta$.

定理 14 每个高斯整数 α 可以表示成素高斯整数的乘积 $\alpha=\pi_1\cdots\pi_n$. 这个表达式实质上是唯一的, 所谓实质上是唯一的是指, 任意其他 α 表示成素高斯整数之积的分解式有相同的因子个数, 并可重新排列使得相应位置的因子是相伴.

为了适当地推广这些概念, 我们首先研究一下高斯整数所满足的不可约多项式方程. 如果 $\alpha=a+b\mathrm{i}$ 是高斯整数, 而不是有理整数, 那么 $b\neq 0$, 并且 α 一定满足一个二次不可约方程. 这就是

$$[x-(a+b\mathrm{i})][x-(a-b\mathrm{i})]=x^2-2ax+(a^2+b^2)=0.$$

它是以有理整数为系数的首一不可约方程. 反过来, 可以证明①, 如果数 $r+s\mathrm{i}$ 在域 $\mathbf{Q}(\mathrm{i})$ 中满足一个整系数首一不可约方程, 那么这个数是高斯整数. 这就给出

定理 15 域 $\mathbf{Q}(\mathrm{i})$ 中的一个数是高斯整数当且仅当它在 $\mathbf{Q}$ 上所满足的首一不可约方程是以整数为系数.

习 题

1. 把下列各高斯整数分解成素因子乘积: $5, 3+\mathrm{i}, 6\mathrm{i}, 11, 1-7\mathrm{i}$.
2. 求出下列每对高斯整数 α_1 和 α_2 的最大公因子, 并把它表示成 $\beta_1\alpha_1+\beta_2\alpha_2$ 的形式:
   (a) $3+6\mathrm{i}$ 和 $12-3\mathrm{i}$, (b) $5+3\mathrm{i}$ 和 $13+18\mathrm{i}$.
3. 求出 13 的所有可能的因子分解 (分解成素高斯整数之积), 并证明: 任意两个分解仅差相伴.
4. 证明: 由高斯整数构成的每个理想都是主理想.
5. (a) 用欧几里得算法证明引理 1. (b) 证明引理 2.
6. 由引理 2 证明定理 14.
7. (a) 证明: 有理素数 p 在 $\mathbf{Z}[\mathrm{i}]$ 中是素的当且仅当方程 $x^2+y^2=p$ 没有整数解 x 和 y.
   (b) 证明: 任意形为 $p=4n+3$ 的有理素数在 $\mathbf{Z}[\mathrm{i}]$ 中是素的.

*8. (a) 证明: 商环 $\mathbf{Z}[x]/(p,x^2+1)$ 既与 $\mathbf{Z}[\mathrm{i}]/(p)$ 同构, 也与 $\mathbf{Z}_p[x]/(x^2+1)$ 同构.
   (b) 证明: $\mathbf{Z}[\mathrm{i}]/(p)$ 是整环当且仅当 p 在 $\mathbf{Z}[\mathrm{i}]$ 中是素的; $\mathbf{Z}_p[x]/(x^2+1)$ 是整环当且仅当 $x^2\equiv -1(\mathrm{mod}\ p)$ 在 $\mathbf{Z}$ 中没有解.
   (c) 假设模 p 乘法群是循环群 (15.3 节定理 6), 证明: 如果 $p=4n+1$, 那么 $x^2\equiv -1(\mathrm{mod}\ p)$ 在 $\mathbf{Z}$ 中有解.

① 在下一节中 (定理 16), 对稍微更一般的情形给出证明.

(d) 证明：$p=4n+1$ 在 $\mathbf{Z}[\mathrm{i}]$ 中不可能是素的.

习题 9 ~ 习题 13 都指的是由数 $a+b\sqrt{-2}$(其中 a 和 b 是整数) 构成的整环 $\mathbf{Z}[\sqrt{-2}]$.

9. 定义范数 $N(a+b\sqrt{-2})=a^2+2b^2$, 并列出它的性质.
10. 证明整环 $\mathbf{Z}[\sqrt{-2}]$ 中的除法算式.
11. 证明整环 $\mathbf{Z}[\sqrt{-2}]$ 中最大公因子的存在性.
12. 叙述并证明 $\mathbf{Z}[\sqrt{-2}]$ 中的唯一因子分解定理.
13. 在 $\mathbf{Z}[\sqrt{-2}]$ 中, 把下列各数因子分解：$5, 1+3\sqrt{-2}, 2+\sqrt{-2}$.
14. (a) 在 $\mathbf{Z}[\sqrt{2}]$ 中求一个不同于 ±1 的单位.

    (b) 证明：在 $\mathbf{Z}[\sqrt{2}]$ 中存在无穷多个不同的单位. (提示：用一个单位的幂.)

14.8 代数整数

一般地, 如果代数数 u 满足的有理数域上首一不可约方程是以整数为系数, 即

$$p(u)=a_0+a_1u+\cdots+a_{n-1}u^{n-1}+u^n=0,\quad a_i \text{ 是整数} \tag{14}$$

其中 $p(x)$ 在 $\mathbf{Q}$ 上是不可约的, 那么称 u 是*代数整数*. 有理数 $\frac{m}{n}$ 所满足的不可约方程刚好是线性方程 $x-\frac{m}{n}=0$. 所以一个有理数是代数整数当且仅当它是一个普通意义下的整数. 于是 $\mathbf{Z}$ 的 (普通) 整数称为*有理整数*, 以便同其他代数整数相区别. 如果代数数 $u\neq0, u$ 和 u^{-1} 都是代数整数, 那么称 u 是*单位*.

在检验给定的代数数是否是代数整数时, 不一定要借助于不可约方程, 而依赖下面的定理.

定理 16 *一个数是代数整数当且仅当它在 $\mathbf{Q}$ 上满足一个整系数首一多项式方程.*

证明 假设 u 是某个整系数首一多项式 $f(x)$ 的根, 在 $\mathbf{Q}$ 上, u 还满足一个不可约多项式 $p(x)$, 它可以取为整系数的. 这些系数的任意公因子可被消去, 因此我们可以假定 $p(x)$ 的系数的最大公因子是 1. 这也就是说, 在 3.9 节的意义下, $p(x)$ 是由所有整系数多项式组成的整环 $\mathbf{Z}[x]$ 中的本原多项式. 因为已知的多项式 $f(x)$ 是首一多项式, 所以也是本原的. 由定理 2 我们知道, 以 u 为根的多项式 $f(x)$ 在 $\mathbf{Q}[x]$ 中一定被 u 所满足的不可约多项式 $p(x)$ 整除, 于是 $f(x)=q(x)p(x)$. 因为 f 和 p 都是本原的, 所以根据 3.9 节引理 3 可以断言, 商式 $q(x)$ 也具有整系数. 那么 $f(x)$ 的首项系数 1 是 q 和 p 的首项系数之积, 因此 $\pm p(x)$ 是首一多项式, 根据定义 (14), 这意味着 u 是代数整数. 证毕

一个数虽然看起来不像代数整数, 但实际上可能是代数整数, 例如, $u=\frac{1+\sqrt{5}}{2}$ 看起来像一个分数, 但它满足方程

$$\left(x-\frac{1+\sqrt{5}}{2}\right)\left(x-\frac{1-\sqrt{5}}{2}\right)=x^2-x-1=0,$$

这个方程是首一整系数方程. 这就暗示我们系统地找出在二次域中是代数整数的那些数. 有理数域 $\mathbf{Q}$ 上的任意二次域 K 可以表示成单代数扩张 $K=\mathbf{Q}(\sqrt{d})$. 不失一般性, 我们可以假定 d 是一个整数且无平方因子 (1 除外). 这就是我们所要考虑的情形:

定理 17 如果 $d\neq 1$ 是无平方因子的整数, 那么在 $d\equiv 2$ 或 $d\equiv 3(\bmod 4)$ 的情形下, $\mathbf{Q}(\sqrt{d})$ 中的代数整数是形为 $a+b\sqrt{d}$(其中系数 a 和 b 是有理整数) 的数. 但是当 $d\equiv 1(\bmod 4)$ 时, $\mathbf{Q}(\sqrt{d})$ 中的代数整数是形为 $a+b\dfrac{1+\sqrt{d}}{2}$(其中 a 和 b 为有理整数) 的数.

证明 作为预备知识, 我们注意到, $a\equiv 1(\bmod 2)$ 意味着 $a=1+2r$, 因此 $a^2=1+4r+4r^2\equiv 1(\bmod 4)$. 换句话说,

$$a\equiv 1(\bmod 2) \quad 可推出\ a^2\equiv 1(\bmod 4), \tag{15}$$

$$a\equiv 0(\bmod 2) \quad 可推出\ a^2\equiv 0(\bmod 4), \tag{16}$$

所以一个平方数总同余于 0 或 1, mod 4.

$\mathbf{Q}(\sqrt{d})$ 中任意数 u 可表示成 $u=\dfrac{a+b\sqrt{d}}{c}$, 其中整数 a,b,c 无公因子. 为了排除有理数的平凡情形, 我们假定 $b\neq 0$. 那么 u 所满足的二次首一不可约方程是

$$\left(x-\frac{a+b\sqrt{d}}{c}\right)\left(x-\frac{a-b\sqrt{d}}{c}\right)=x^2-\frac{2a}{c}x+\frac{a^2-db^2}{c^2}=0. \tag{17}$$

如果 u 是代数整数, 那么这些系数 $\dfrac{2a}{c}$ 和 $\dfrac{a^2-db^2}{c^2}$ 也一定是整数. 所以 $\dfrac{4a^2}{c^2}, \dfrac{4a^2-4db^2}{c^2}$ 和 $\dfrac{4db^2}{c^2}$ 都必是整数, 所以 $c|2a, c^2|4db^2$. 因为已假定 d 不含平方因子, 所以包含在 c 中的任意素数 $p\neq 2$ 一定同时整除 a 和 b^2, 这与 a,b,c 无公因子 (± 1 除外) 的假定相矛盾. 由于类似的理由, $4|c$ 是不可能的, 所以只能选取 $c=1$ 和 $c=2$.

现在考虑 $d\equiv 2$ 或 $d\equiv 3(\bmod 4)$ 的情形, 取 $c=2$. 在这种情形下, (17) 式最后的系数 $\dfrac{a^2-db^2}{4}$ 一定是整数, 于是 $a^2\equiv db^2(\bmod 4)$. 如果 $b\equiv 1(\bmod 2)$, 则 $b^2\equiv 1(\bmod 4)$, 并且 $a^2\equiv db^2\equiv 2$ 或 $3(\bmod 4)$. 这与法则 (15) 和 (16) 相矛盾. 如果 $b\equiv 0(\bmod 2)$, 则 $a^2\equiv 0(\bmod 4)$, 因而 $a\equiv 0(\bmod 2)$, 所以 a,b,c 有公因子 2. 无论哪一种情形, 我们都得出 $c=1$, 所以 $\mathbf{Q}(\sqrt{d})$ 的所有代数整数是形为 $a+b\sqrt{d}$ 的数. 反过来, 这种形式的数所满足的首一方程 (17) 具有整系数.

剩下的情形 $d\equiv 1(\bmod 4)$ 可以类似处理, 除了可能出现 $a\equiv b\equiv 1(\bmod 2)$ 的情形以外, 其余都类似.

推论 $\mathbf{Q}$ 上任意二次域中, 所有代数整数组成的集合是一个整环.

证明 定理 17 中所表示的代数整数的和、差、积仍然是这种形式的代数整数.

证毕

下面的任务就是把这个推论推广到任意代数数域上.

习 题

1. 证明：每个单位根是代数整数.
2. (a) 求 $\mathbf{Q}(w)$ 中的所有代数整数和所有单位, 这里 w 是复三次单位根.
   (b) 证明：$\mathbf{Q}(w)$ 中每个单位都是单位根.
3. 完成定理 17 的第二种情形 ($d \equiv 1(\mathrm{mod}\ 4)$) 的证明.
4. (a) 证明：任意代数数可以写成商 $\dfrac{u}{b}$, 这里 u 是代数整数, b 是有理整数 (即 $\mathbf{Z}$ 中的整数).
   (b) 证明：代数数组成的任意域 K 是由 K 中的所有代数整数组成的整环的商域.

*5. 求出 $\mathbf{Q}(\sqrt{2}, \mathrm{i})$ 中的所有代数整数.

14.9 代数整数的和与积

这一节来证明下述结果：

定理 18 *所有代数整数的集合是一个整环.*

下面是定理的直接推论.

推论 *在由代数数组成的任意域 K 中, 全体代数整数构成一个整环.*

定理 18 的一个启发性的证明依赖于对代数整数生成的加法群的分析. 如果 $v_1, \cdots, v_n$ 是任意代数数, 我们令 $G = [v_1, \cdots, v_n]$ 表示在由全体复数组成的加法群中由这些代数数生成的子群①. 这个群只是由可表示成形式

$$u = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n \quad (a_i \text{ 是有理整数}) \tag{18}$$

的所有数组成. 回忆一下, 在由 v 生成的加法循环子群中, 自然倍数 $av = a \times v$ 就是 v 的"幂".

引理 1 *群 $G = [v_1, \cdots, v_n]$ 的任意子群 S 也可以由 n 个或更少一些数生成.*

证明 对每个下标 k, 设 G_k 是由 G 的后 $n-k+1$ 个生成元生成的子群 $[v_k, \cdots, v_n]$, 所以 G_k 是由所有形为 $a_k v_k + \cdots + a_n v_n$ 的和构成. 在 G_k 的位于给定子群 S 的元素中间选取一个元素

$$w_k = c_k v_k + c_{k+1} v_{k+1} + \cdots + c_n v_n, \tag{19}$$

使其中第一个系数 c_k 有最小的正值 (这是可能的). (如果对每个元素, v_k 的系数都是零, 则令 $w_k = 0$.) 如果 $w = b_k v_k + \cdots$ 是 G_k 中 S 的任意其他元素, 它的

① 这个加法群有时称为 $\mathbf{Z}$- 模, 因为它的元素可以用 $\mathbf{Z}$ 中的标量来乘.

第一个系数 b_k 可以写成 $b_k = q_k c_k + r_k$, 具有一个非负余数 $r_k < c_k$. 那么这个差 $w - q_k w_k = r_k v_k + \cdots$ 就在群 G_k 和 S 中, 并有非负的第一个系数 r_k, 它小于最小的正值 c_k, 因此 $r_k = 0$, 于是 G_k 中任意 S 的元素 w 给出 G_{k+1} 中的一个元素 $w' = w - q_k w_k$.

这样选取的 n 个元素 $w_1, \cdots, w_n$ 生成整个群 S, 这是因为对 S 中任意给定元素 w, 我们就可以找到 q_1, 使 $w - q_1 w_1$ 只依赖于 $v_2, \cdots, v_n$, 然后又可找到某个 q_2, 使 $w - q_1 w_1 - q_2 w_2$ 只依赖于 $v_3, \cdots, v_n$, 等等, 最后有 $w = \Sigma q_i w_i$. 　　证毕

引理 2　数 u 是代数整数当且仅当由 u 的所有幂 $1, u, u^2, u^3, \cdots$ 生成的加法群可以由有限个元素生成.

证明　如果 u 是代数整数, 则它满足整系数的 n 次首一方程 (14). 这个方程把 u^n 表示成群 $G = [1, u, \cdots, u^{n-1}]$ 中的一个元素, 这个群是由 u 的低于 n 次的幂生成的. 通过迭代, 上述同一个方程可以用来把 u 的任意高次幂表示成这个群中的一个元素. 所以 u 满足引理 2 的准则.

反过来, 假定由 $1, u, u^2, \cdots$ 生成的群 G 可以由 G 的任意 n 个数 $v_1, \cdots, v_n$ 生成. u 与 G 的任意元素 $\Sigma a_j u^j$ 的乘积仍然是 G 的元素 $\Sigma a_j u^{j+1}$, 所以每个乘积 uv_i 一定在 G 中, 于是它一定可以按照生成元表示成 $uv_i = \sum_j a_{ij} v_j$, 其中 a_{ij} 是整数. 这些表达式给出如下形式的 n 个 v 的齐次方程

$$
\begin{aligned}
(a_{11} - u)v_1 + a_{12}v_2 + \cdots + a_{1n}v_n &= 0, \\
a_{21}v_1 + (a_{22} - u)v_2 + \cdots + a_{2n}v_n &= 0, \\
&\vdots \\
a_{n1}v_1 + a_{n2}v_2 + \cdots + (a_{nn} - u)v_n &= 0.
\end{aligned}
$$

这组方程有一组不全为零的解 $v_1, v_2, \cdots, v_n$, 所以系数矩阵的行向量一定是线性相关的 (7.7 节定理 13 的推论). 这个系数矩阵可以写成 $\boldsymbol{A} - u\boldsymbol{I}$, 这里 $\boldsymbol{A} = (a_{ij})$. 因为它是奇异的, 所以它的行列式是零, 于是

$$
|\boldsymbol{A} - u\boldsymbol{I}| = (-1)^n u^n + b_{n-1}u^{n-1} + \cdots + b_n = 0, \tag{20}
$$

这里系数 b_i 是整数 a_{ij} 的某一多项式, 因而它们都是整数. 这个方程 (20) 意味着[①] u 是代数整数, 正如引理中所要求的.

引理 2 的结论可以重述如下:

推论　如果代数数 u 的所有正次幂都在由一组有限个数 $y_1, \cdots, y_n$ 生成的加法群中, 那么 u 是代数整数.

---

① 注意, 根据第 10 章的意义, (20) 式就是 $\boldsymbol{A}$ 的特征多项式.

证明 由 $1, u, u^2, \cdots$ 生成的群 S 是由 $1, y_1, \cdots, y_n$ 生成的群的一个子群. 因此根据引理 1, 这个子群 S 可以由它的有限多个元素生成, 因此根据引理 2, 数 u 是代数整数. 证毕

现在回来证明定理 18. 如果 u 和 v 是代数整数, 我们应指出 $u+v$ 和 uv 都是代数整数. 这个假设条件意味着所有的幂 u^k 和 v^k 分别可以按照有限多个幂 $1, u, \cdots, u^{n-1}$ 和 $1, v, \cdots, v^{r-1}$ 来表示. 所以每个幂 $(uv)^k = u^k v^k$ 和 $(u+v)^k$ 在由乘积 $1, u, uv, uv^2, \cdots, u^{n-1}v^{r-1}$ 生成的加法群中. 根据推论得出, uv 和 $u+v$ 是代数整数, 这正是定理所要求的.

习 题

1. 通过列出适当的整系数首一方程, 明显地证明下列各数都是代数整数:
   (a) $\sqrt{2}+\sqrt{3}$, (b) $\mathrm{i}+\omega$, (c) $\sqrt{7}+\dfrac{1+\sqrt{5}}{2}$.
2. (a) 证明: 如果数 $v_1, \cdots, v_n$ 在 $\mathbf{Q}$ 上线性无关, 那么在 $G=[v_1, \cdots, v_n]$ 中的任意有有限指数的子群 S 也可以由 n 个线性无关的数 $w_1, \cdots, w_n$ 生成.
   (b) 证明: 任意这样的子群 S 与整个群 G(群) 同构.
3. 如果数 $v_1, \cdots, v_n$ 在 $\mathbf{Q}$ 上线性无关, 指出怎样用引理 1 中对于 $G=[v_1, \cdots, v_n]$ 的子群 S 求出的那组基来计算 S 在 G 中的指数. (提示: 求出 S 的每个陪集的代表元.)

*4. 证明: 群 $G=[v_1, \cdots, v_n]$ 没有不同子群的无限升链. 也就是证明: 给出子群的无穷序列 $S_1 \leqslant S_2 \leqslant S_3 \leqslant \cdots \leqslant G$, 则存在一个下标 m, 使得 $S_m = S_{m+1} = S_{m+2} = \cdots$. (提示: 把引理 1 应用到全体群 S_k 的并.)

5. (a) 证明: 包含在普通整数环 $\mathbf{Z}$ 中的每个模是 $\mathbf{Z}$ 的一个理想.
   (b) 列出一个包含在高斯整数的整环 $\mathbf{Z}[\mathrm{i}]$ 中不是 $\mathbf{Z}[\mathrm{i}]$ 的理想的模.

*6. 证明: 如果代数数 u 满足一个首一多项式方程, 这个方程的其他系数都是代数整数, 那么 u 也是代数整数.

14.10 二次代数整数的因子分解

为了说明代数整数因子分解理论, 我们更详细地考虑最简单的情形, 即二次代数整数的因子分解. 也就是说, 我们考虑 $\mathbf{Q}(\sqrt{d})$ 的代数整数 (像定理 17 中所描述的那种代数整数) 的因子分解. 为此目的所用的基本工具是范数概念.

范数的公式依赖于域, 但是范数的概念在所有情况下, 甚至对于高次代数数域, 都是一样的. 本质上, 范数是通过域的自同构来定义的. 根据定理 6, 二次域 $\mathbf{Q}(\sqrt{d})$ 有一个自同构 $u=a+b\sqrt{d} \longleftrightarrow \bar{u}=a-b\sqrt{d}$, 它把每个数映射到它的共轭数 $\bar{u}$.

定义 $\mathbf{Q}(\sqrt{d})$ 的数 $u=a+b\sqrt{d}$ 的范数 $N(u)$ 是 u 和它的共轭 $\bar{u}$ 的乘积 $u\bar{u}$:

$$N(u) = u\bar{u} = (a+b\sqrt{d})(a-b\sqrt{d}). \tag{21}$$

因为对应 $u \longleftrightarrow \bar{u}$ 是同构, $\overline{uv} = \bar{u} \cdot \bar{v}$, 所以

$$N(uv) = N(u)N(v). \tag{22}$$

于是范数把这个域中的代数整数的任意分解 $w = uv$ 转换成有理整数 $N(w)$ 的分解 $N(w) = N(u)N(v)$. (代数整数的范数是有理整数, 见习题 1.)

范数的性质基本上依赖于 d 是正的还是负的, 也就是依赖于 $\mathbf{Q}(\sqrt{d})$ 是实二次域还是复二次域. 如果 $d < 0$, 则 $N(u)$ 就是 $|u|^2$, 即 u 的绝对值平方, 除了 $u = 0$ 外, 它是正的. 如果 $d > 0$, 则 $N(u) = a^2 - b^2 d$ 可以是正的也可以是负的. 这个差别出现在 $\mathbf{Q}(\sqrt{d})$ 的单位的群 U 中, 正如我们将看到的.

引理 1 代数整数 $u \in \mathbf{Q}(\sqrt{d})$ 是单位当且仅当 $N(u) = \pm 1$.

证明 显然, $N(1) = 1$; 此外, $N(u)$ 一定是有理整数. 因此, 如果对某一个其他代数整数 $v \in \mathbf{Q}(\sqrt{d})$ 有 $uv = 1$, 那么有 $N(u)\ N(v) = N(uv) = 1$, 因此 $N(u) = \pm 1$. 反过来, 如果 $N(u) = u\bar{u} = \pm 1$, 那么 $u(\pm \bar{u}) = 1$, 于是 u 是 $\mathbf{Q}(\sqrt{d})$ 的单位.

类似的论证可应用到一般代数数域上.

把引理 1 和定理 17 结合起来, 我们可以确定任意复二次域 $\mathbf{Q}(\sqrt{-d})$($d > 0$, 无平方因子的整数) 的全部单位, 那么 $\mathbf{Q}(\sqrt{-d})$ 的代数整数具有形式 $u = m + n\alpha (m, n \in \mathbf{Z})$, 这里

$$\alpha = \begin{cases} \sqrt{-d}, & \text{当 } d \not\equiv 3 \pmod 4 \\ \dfrac{1 + \sqrt{-d}}{2}, & \text{当 } d \equiv 3 \pmod 4 \end{cases}$$

相应地, u 的范数满足

$$N(u) = \begin{cases} m^2 + n^2 d, & \text{当 } d \not\equiv 3 \pmod 4 \\ \left(m + \dfrac{n}{2}\right)^2 + \dfrac{n^2 d}{4}, & \text{当 } d \equiv 3 \pmod 4 \end{cases}$$

当 $d \not\equiv 3 (\bmod\ 4)$ 且 $d > 1$ 时, $m^2 + n^2 d \leqslant 1$ 只有当 $m = \pm 1, n = 0$ 时才有可能. 同样, 如果 $d \equiv 3 (\bmod\ 4)$ 且 $d > 3$, 那么 $d \geqslant 7$, 并且 $N(u) \geqslant \dfrac{7n^2}{4} > 1$, 除非 $n = 0$. 因此又一次说明 $\mathbf{Q}(\sqrt{-d})$ 的单位只能是 ± 1. 这就证明了

定理 19 存在不同于 ± 1 的单位的复二次域只能是 $\mathbf{Q}(\sqrt{-1})$ 和 $\mathbf{Q}(\sqrt{-3})$.

$\mathbf{Q}(\sqrt{-1})$ 的单位是 ± 1 和 $\pm \mathrm{i}$; $\mathbf{Q}(\sqrt{-3})$ 的单位是 $\omega = \dfrac{1 + \sqrt{-3}}{2}$ 的各次幂, ω 是六次本原单位根.

实二次域有无穷多个单位. 例如, $1 + \sqrt{2}$ 是 $\mathbf{Q}(\sqrt{2})$ 的单位, 因为 $N(1 + \sqrt{2}) = -1$. 因此 $1 + \sqrt{2}$ 的所有次幂 $(1 + \sqrt{2})^{\pm k}$ 都是单位.

虽然对于很多二次代数整数的环, 分解成素因子的因子分解是唯一的, 但在 $\mathbf{Q}(\sqrt{-5})$ 中情况并非如此. 例如, 考虑数 6 的因子分解:

$$6 = 2 \times 3 = (1+\sqrt{-5}) \times (1-\sqrt{-5}). \tag{23}$$

如果 $\mathbf{Q}(\sqrt{-5})$ 的两个代数整数 u 和 v 满足 $uv = 6$, 那么 $N(u)N(v) = N(6) = 36$. 于是 6 的真因子 u 的范数将是 $2^2 3^2$ 的真因子, 所以只有 $N(u) = 2, 3, 4, 6, 9, 12, 18$ 几种情况需要研究. 因为在这些情况中, $N(v)$ 分别为 $18, 12, 9, 6, 4, 3, 2$, 所以只须考虑 $u = 2, 3, 4, 6$ 的情形. 由 $N(m+n\sqrt{-5}) = m^2 + 5n^2$ 容易看出, 所有可能的因子都已列在 (23) 式中.

在上述例子中, 我们可以考虑用理想的乘积 (如 13.4 节中所述) 代替数的乘积, 以此来补救唯一因子分解定理. 我们发现主理想 (2), (3), $(1+\sqrt{-5})$ 和 $(1-\sqrt{-5})$都不是素理想. 相关的素理想是 $P = (2, 1+\sqrt{-5}), Q = (3, 1+\sqrt{-5})$, 这是用它们在 $\mathbf{Z}(\sqrt{-5})$ 中的基来描述的. 这些理想不是主理想, 把它们平方起来:

$$P^2 = (4, 2+2\sqrt{-5}, 6) = (2)$$

$$Q^2 = (9, 3+3\sqrt{-5}, 6) = (3)$$

这表明 (2) 和 (3) 不是素理想.

为证明 P 是 $\mathbf{Z}[\sqrt{-5}]$ 中的素理想, 我们注意, $(m+n\sqrt{-5}) \in P$ 当且仅当 $m+n \equiv 0 (\bmod 2)$. 因此 $\mathbf{Z}[\sqrt{-5}]/P$ 只包含两个元素, 它是域 $\mathbf{Z}_2$. 因此根据 13.3 节定理 6, P 是素理想. 类似地, $\mathbf{Z}[\sqrt{-5}]/Q$ 是 $\mathbf{Z}_3$, 所以 Q 也是素理想.

总之, 我们证明了 $\mathbf{Z}[\sqrt{-5}]$ 的理想 (6) 具有分解成素理想的唯一因子分解 $(6)= P^2Q^2$.

我们在整环 $\mathbf{Z}[\sqrt{-5}]$ 中推导出的这种理想的唯一分解仅仅是用以说明, 理想的概念怎样可以系统地用来建立代数整数环上的唯一因子分解定理, 而通常的因子分解在这个整环上是不唯一的. 通过进一步推理, 我们可以建立"理想论基本定理": *在代数数域 K 中由所有代数整数组成的整环 D 中, 每个理想除次序外可以唯一地表示成素理想的乘积. 特别是, 整环的每个代数整数 u 确定一个主理想 (u), 在上述意义下, 它有唯一的因子分解.*

习 题

1. (a) 证明: 在任意二次域中, 代数整数的范数是有理整数.
   (b) 证明: 如果 $u = a + b\sqrt{d}$ 不是有理数, 那么 $N(u)$ 是 u 所满足的首一不可约多项式方程的常数项.

2. 求 $\mathbf{Q}[\sqrt{-7}]$ 中的全部单位.
3. 证明：二次域 $\mathbf{Q}(\sqrt{-d})$(其中 $d>0$) 中单位的个数是有限的, 并证明：每个单位是单位根.
*4. 证明：任意给定的代数数域中的全体单位根构成循环群.
5. 叙述并证明 $\mathbf{Z}[\omega]$ 的除法算式, 这里 $\omega=\dfrac{-1+\sqrt{-3}}{2}$. (提示：任意 β 的整数倍数把复平面分割成无穷多个等边三角形.)
*6. 设 D 是任意整环, 在这个整环中范数 $N(\alpha)$ 定义如下：(i) 当 $a\neq 0, N(\alpha)$ 是正整数; (ii) $N(\alpha\beta)=N(\alpha)N(\beta)$; (iii) 给定 α 和 $\beta\neq 0$, 存在 γ 和 ζ 使得 $\alpha=\beta\gamma+\zeta, N(\zeta)<N(\beta)$.
   (a) 证明：D 是唯一因子分解整环.
   (b) 证明：D 中每个理想是主理想.

第 15 章　伽罗瓦理论

15.1　方程的根域

历史上很多代数学家试图用明显的公式求解实多项式方程和 (以后的) 复多项式方程. 在他们的努力下, 求出了一般的二次、三次和四次方程的“根式解”, 这些求解公式我们已在第 5 章推导过. 但是对于五次方程, 多次想得到类似的求解公式, 结果都失败了.

这是为什么? 其原因最后被伽罗瓦 (Evariste Galois) 发现, 他指出一个方程有根式解当且仅当与它相联系的自同构群在纯群论意义下是“可解”的. 这里所说的自同构是指由这个方程的所有根生成的扩域的, 使方程的所有系数保持固定的那些自同构. 在最后这一章里, 我们从讨论给定域 F 上的已知多项式 $p(x)$ 的所有根生成的扩域开始, 按照现代的形式介绍伽罗瓦最本质的论证. 这个域就是所谓 $p(x)$ 的“根域”,[①]现在我们正式地给出它的定义.

定义　F 的扩张 N 如果满足下列条件则称为系数在 F 中的 $n(\geqslant 1)$ 次多项式 $f(x)$ 的根域: (i) $f(x)$ 在 N 中可以分解成线性因子 $f(x) = c(x-u_1)\cdots(x-u_n)$; (ii) N 是 F 上添加 $f(x)$ 的全部根而生成的, 即 $N = F(u_1,\cdots,u_n)$.

如果 $f(x) = ax^2+bx+c(a\neq 0)$ 是 F 上的二次多项式, 它有共轭根[②] $u_j = \dfrac{-b\pm\sqrt{b^2-4ac}}{2a}, j=1,2$, 由 $f(x)=0$ 的一个根 u_1 生成的 F 的单扩张 $K=F(u_1)\cong F[x]/(f(x))$ 已经是 f 在 F 上的根域. 这是因为 $u_2 = \dfrac{c}{au_1}$, 因此 $f(x)$ 在域 $K=F(u_1)$ 中可以分解成线性因子 $f(x)=a(x-u_1)(x-u_2)$.

可是, 对于三次不可约多项式, 这个结论一般来说是不正确的. 例如, $\mathbf{Q}$ 上不可约多项式 x^3-5 的根域 N 是 $\mathbf{Q}(\sqrt[3]{5},\omega\sqrt[3]{5},\omega^2\sqrt[3]{5}) = \mathbf{Q}(\sqrt[3]{5},\omega)$,其中 $\omega = \dfrac{-1+\mathrm{i}\sqrt{3}}{2}$ 是复三次单位根. 由 5 的实三次根生成的有理数域的实扩张 $\mathbf{Q}(\sqrt[3]{5})\cong\mathbf{Q}[x]/(x^3-5)$ 在 $\mathbf{Q}$ 上的次数是 3, 而包含 5 的所有三次根的 $\mathbf{Q}$ 的最小扩张是 $N=\mathbf{Q}(\sqrt[3]{5},\omega)$. 因为 ω 满足分圆方程 $\omega^2+\omega+1=0$, 所以域 N 在 $\mathbf{Q}(\sqrt[3]{5})$ 上的次数是 2. 当我们把 x^3-5 的根域 N 看作 $\mathbf{Q}$ 上的一个向量空间时, 它就有基 $(1, \sqrt[3]{5}, \sqrt[3]{25}, \omega, \omega\sqrt[3]{5}, \omega\sqrt[3]{25})$, 于

① 有的代数书中称这个域为分列域. —— 译者注

② 当然, 多项式 $f(x)$ 的根是指满足 $f(x)=0$ 的数 x, 这个 x 也称为 $f(x)$ 的零点.

是它是 $\mathbf{Q}$ 的六次扩张.

可以用已知的单代数扩张的存在性得到一般的关于根域存在性命题, 如下所述:

定理 1 任意域上的任意多项式都有一个根域.

对于一次多项式, 其根域刚好是基域 F; 因此我们可以对 $f(x)$ 的次数 n 用归纳法. 假定这个定理对所有域 F 和所有 $n-1$ 次多项式都成立, 并设 $p(x)$ 是已知多项式 $f(x)$ 的在 F 上一个不可约因子. 根据 14.3 节的定理 5, 存在一个由 $p(x)$ 的根 u 生成的单扩张 $K=F(u)$. 在 K 上, $f(x)$ 有根 u, 因而有因子 $x-u$, 所以 $f(x)=(x-u)g(x)$, 商 $g(x)$ 是 K 上 $n-1$ 次多项式, 根据归纳法假定, 由 $g(x)$ 的 $n-1$ 个根生成 K 上的一个根域 N. 这个域 N 也是 $f(x)$ 的根域.

我们将在下一节 (定理 2) 证明, 给定基域 F 上的已知多项式 $f(x)$ 的所有根域是同构的, 所以说它是 $f(x)$ 在 F 上的根域是合理的.

附录 定理 1 可以用来 (纯代数地) 构造任意有限域或可数域 F 的代数完全扩张, 如下所述. 域 F 上的 n 次多项式的个数是有限的或可数的, 如果 F 是可数的, 那么多项式个数是 $\mathbf{d}^{n+1}=\mathbf{d}(\mathbf{d}=$ 无限可数). 因此 F 上所有多项式的个数是可数的 (参看 12.2 节习题 14), 于是我们可以把这些多项式排成序列 $p_1(x),p_2(x),p_3(x),\cdots$.

现在设 F_1 是 $p_1(x)$ 在 F 上的根域, F_2 是 $p_2(x)$ 在 F_1 上的根域, $\cdots$; 一般地, 设 F_n 是 $p_n(x)$ 在 F_{n-1} 上的根域. 最后, 设 F^* 是由出现在其中一个 F_n 中, 因而出现在 F_n 的所有后继中的所有元素组成的集合. 如果 a 和 b 是 F^* 的任意两个元素, 那么它们一定都在某一个 F_n 中, 因而都在 F_n 的所有后继中, 所以 $a+b,ab$ 和 $\dfrac{a}{b}(b\neq 0)$ 在 F_n 和它的所有后继中也一定有相同的值, 这就表明 F^* 是一个域.

为了证明 F^* 是代数完全域, 我们设 $g(x)$ 是 F^* 上任意多项式, g(x) 的全部系数都在某一个 F_n 中, 因而这些系数在 F 上是代数的. 那么利用 14.5 节定理 9, 我们可以求出 $g(x)$ 的一个非零倍式 $h(x)$, 其系数在 F 中 (见下面的习题 5). 但是对 $h(x)$, 有一个适当的 $F_{m-1},h(x)$ 在 F_{m-1} 上的根域是 $F_m,h(x)$ 在 F_m 中一定可以分解成线性因子, 因此 $h(x)$ 的因子 $g(x)$ 在 F_m 中同样也可分解成线性因子. 所以 $g(x)$ 在较大的域 F^* 上也可以分解成线性因子, 因此 F^* 是代数完全域, 进一步有, F^* 的每个元素在 F 上是代数的.

用一般的良序集合和所谓超限归纳法来代替序列, 可把上述推理过程加以修改[①], 以便应用到任意域 F 上, 这种修改建立了下面关于代数基本定理的重要的部分的推广: 任意域 F 都有一个代数完全扩张.

① 详细证明见 B. L. van der Waerden, *Moderne Algebra,* Part I, Berlin, 1930. (中译本: B. L. 范德瓦尔登, 代数学 I, 丁石孙等译, 科学出版社, 1965.)

15.2 唯一性定理

我们现在来证明定理 1 所述的根域的唯一性 (精确到同构).

定理 2 域 F 上已知多项式 $f(x)$ 的任意两个根域 N 和 N' 同构, N 到 N' 的同构可以如此选择, 使得 F 的元素保持固定.

证明 "根域是唯一的"这个断言实质上是"同一个不可约多项式的两个不同根生成同构的单代数扩张"这一事实的直接推论 (14.3 节定理 6). 特别是, 不可约多项式 $p(x)$ 的两个根域 $N = F(u_1, \cdots, u_n)$ 和 $N' = F(u_1', \cdots, u_n')$ 分别包含由 $p(x)$ 的根 u_1 和 u_1' 生成的同构单扩张 $F(u_1)$ 和 $F(u_1')$. 因此存在 $F(u_1)$ 到 $F(u_1')$ 的同构 T. 现在只需要把这个同构适当地扩张到整个根域上. 这种扩张的基本方法由下面引理给出.

引理 1 如果域 F 和 F' 之间的同构 S 把不可约多项式 $p(x)$ 的系数映射到 F' 上多项式 $p'(x)$ 相应的系数, 并设 $F(u)$ 和 $F'(u')$ 分别是由这两个多项式的根 u 和 u' 生成的单扩张, 那么 S 可以扩张成 $F(u)$ 到 $F'(u')$ 的同构 S^*, 在这个同构之下, $uS^* = u'$.

证明 恰好同 14.3 节定理 6 的讨论一样, 我们所需要的同构 S^* 由公式

$$(a_0 + a_1u + \cdots + a_{n-1}u^{n-1})S^* = a_0S + (a_1S)u' + \cdots + (a_{n-1}S)(u')^{n-1} \qquad (1)$$

明显地给出, 其中所有 $a_i \in F, n$ 是 u 在 F 上的次数.

引理 2 如果 F 到 F' 的同构 S 把多项式 $f(x)$ 映射到 $f'(x)$, 并设 $N \supset F$ 和 $N' \supset F'$ 分别是 $f(x)$ 和 $f'(x)$ 的根域, 那么同构 S 可以扩张成 N 到 N' 的同构.

通过对次数 $m = [N : F]$ 用归纳法可以证明这个引理, 当 $m = 1$ 时, 这是显然的, 因为这时 S 已经扩张到 N; 因此取 $m > 1$, 并且假定引理对于某域 F 上次数小于 m 的所有根域 N 都正确. 因为 $m > 1, f(x)$ 的根不全都在 F 中, 所以 $f(x)$ 中至少有一个不可约因子 $p(x)$, 它的次数为 $d > 1$. 设 u 是 $p(x)$ 在 N 中的根, 而在给定的同构 S 之下, $p'(x)$ 是 $f'(x)$ 的对应于 $p(x)$ 的因子. 那么根域 N' 包含 $p'(x)$ 的根 u', 并根据引理 1, 给定的同构 S 可以扩张成同构 S^*, 满足

$$uS^* = u', \quad [F(u)]S^* = F'(u'), \quad p(u) = 0, \quad p'(u') = 0. \qquad (2)$$

因为 N 是在 F 上添加 $f(x)$ 的全部根而生成, N 一定是在较大的域 $F(u)$ 上添加这些根而生成, 所以 N 是 $f(x)$ 在 $F(u)$ 上的根域, 其次数是 $\frac{m}{d}$. 根据同样的理由, N' 是 $f'(x)$ 在 $F'(u')$ 上的根域. 因为 $\frac{m}{d} < m$, 所以根据引理的归纳假定可以断言, (2) 表示的同构 S^* 可以从 $F(u)$ 扩张到 N. 这就证明了引理 2.

两个根域 N 和 N' 都是同一个基域 F 的扩张, 并且 S 是 F 到自身上的恒等映射, 这种情形下, 引理 2 表明 N 与 N' 同构, 因此也就证明了定理 2.

习　题

1. 求下列多项式在 $\mathbf{Q}$ 上的根域的次数:

   (a) $x^3-x^2-x-2=0$, (b) $x^3-2=0$,

   (c) $x^4-7=0$, (d) $(x^2-2)(x^2-5)=0$.

2. 证明: 域 F 上 n 次多项式的根域在 F 上的次数至多是 $n!$.
3. (a) 证明: 如果 ξ 是 n 次本原单位根, 那么 $\mathbf{Q}(\xi)$ 是 $x^n-1=0$ 在 $\mathbf{Q}$ 上的根域.

   (b) 当 $n=3,4,5,6$ 时, 计算根域的次数.

4. 证明: 特征为 p 的任意代数完全域包含一个子域与 15.1 节附录中所构造的域同构.

*5. 设 $g(x)=a_0+a_1x+\cdots+a_nx^n$ 的系数在域 F 上是代数的, 证明: $g(x)$ 是系数在 F 中的某非零多项式 $h(x)$ 的因子. (*提示*: 构造 $g(x)$ 在 $F(a_0,\cdots,a_n)$ 上的根域; 在这个根域上把 $g(x)$ 分解成线性因子 $x-u_i$; u_i 在 F 上是代数的, 并满足不可约多项式方程 $h_i(x)$; 令 $h(x)=\Pi h_i(x)$.)

6. 设 $p\in\mathbf{Q}[x]$ 是任意有理系数首一多项式, 并设 $z_1,\cdots,z_n$ 是它的复根. 证明: $\mathbf{Q}(z_1,\cdots,z_n)$ 是 p 在 $\mathbf{Q}$ 上的根域.

15.3 有 限 域

系统地运用根域的性质, 我们可以得到所有有限域 (含有有限个元素的域) 的一个完整论述. 因为特征为 ∞ 的域总包含一个与有理数域同构的无限子域 (13.8 节定理 14), 所以每个有限域具有一个素特征 p. 不失一般性, 我们可以假定 F 包含模 p 整数域 $\mathbf{Z}_p$(见 13.7 节定理 12 的推论). 那么有限域 F 是 $\mathbf{Z}_p$ 的有限扩张, 所以在 $\mathbf{Z}_p$ 上有一组基 $u_1,\cdots,u_n$. F 中每个元素可唯一地表示成线性组合 Σa_iu_i. 这里每个系数恰好可按 p 种方法在 $\mathbf{Z}_p$ 中取值, 所以 F 中总共有 p^n 个元素. 这就证明了:

定理 3 *有限域中元素的个数 q 是它的特征的幂 p^n.*

在含有 $q=p^n$ 个元素的有限域 F 中, 全体非零元素构成 $q-1$ 阶乘法群. 那么这个群中每个元素的阶是 $q-1$ 的一个因子. 所以 F 的每个元素满足方程 $x^{q-1}=1$. 因此 F 的所有元素 $a_1,a_2,\cdots,a_q$(包含着零) 满足方程

$$x^q-x=0,\quad q=p^n. \tag{3}$$

因此乘积 $(x-a_1)(x-a_2)\cdots(x-a_q)$ 是 x^q-x 的因子, 乘积中的这些因子是互素多项式, 每一个都整除 x^q-x. 因为这个乘积同 x^q-x 一样, 也是 q 次的首一多项式, 所以我们得到

$$x^q - x = (x - a_1)(x - a_2)\cdots(x - a_q). \tag{4}$$

因此 F 是 $x^q - x$ 在 $\mathbf{Z}_p$ 上的根域. 元素个数相同的任意其他有限域 F' 是同一个方程的根域, 因此根据根域的唯一性定理 (定理 2), 它与 F 同构. 这个推理证明了

定理 4 *元素个数相同的任意两个有限域同构.*

下面考虑这样一个问题: 实际上存在哪些有限域? 为了列出一个有限域, 我们自然要构造多项式 $x^q - x$ 在 $\mathbf{Z}_p$ 上的根域 N. 我们现在证明, 所要求的根域恰好由这个多项式的全部根组成.

引理 *多项式 $x^q - x$ 在它的根域 N 中有 q 个不同的线性因子.*

用反证法来证明. 如果 $x^q - x$ 有一个重因子 $x - u$, 我们可以写成 $x^q - x = (x-u)^2 g(x)$. 比较它们的形式导数 (3.1 节习题 7), 我们将有

$$(x^q - x)' = qx^{q-1} - 1 = -1,$$

$$[(x-u)^2 g(x)]' = (x-u)[2g(x) + (x-u)g'(x)],$$

因此 $x - u$ 是 -1 的因子, 得出矛盾. 这就证明了引理.

另一方面, $x^q - x$ 的根 $u_1, \cdots, u_q$ 中任意两个根之和是一个根, 这是因为, 在特征为 p 的任意域中有 $(a \pm b)^p = a^p \pm b^p$, 所以如果 $a^{p^n} = a$ 和 $b^{p^n} = b$, 那么

$$(a \pm b)^{p^n} = a^{p^n} \pm b^{p^n} = a \pm b.$$

两个根之积 ab 也是一个根, 这因为 $(ab)^{p^n} = a^{p^n} b^{p^n} = ab$. 还有对两个根之商, 类似的结果成立. 所以 $x^q - x$ 的所有 q 个根组成的集合是根域 N 的一个子域; 因为这个子域包含所有的根, 所以它实际上一定是整个根域 N. 这就意味着, 我们已经构造出含有 q 个元素的域, 因此有

定理 5 *对任意素数 p 和任意正整数 n, 存在一个含有 $p^n = q$ 个元素的有限域: $x^q - x = 0$ 在 $\mathbf{Z}_p$ 上的根域.*

根据定理 4 和定理 5, 存在一个且本质上只存在一个含有 p^n 个元素的域. 这个域有时称为*伽罗瓦域* $GF(p^n)$. 这个域的乘法群的结构可以完整地描述如下.

定理 6 *在任意有限域 F 中, 所有非零元素组成的乘法群是循环群.*

证明 F 中每个非零元素是 $q-1$ 次单位根, 即它满足方程 $x^{q-1} = 1$, 这里 q 是 F 中元素的个数. 为了证明这个群是循环群, 我们必须求出 F 中 $q-1$ 次本原单位根, 它的低于 $q-1$ 次的幂都不等于 1; 那么本原单位根的所有次幂跑遍整个群. 为此, 把 $q-1$ 写成不同素数幂的乘积

$$q - 1 = p_1^{e_1} p_2^{e_2} \cdots p_r^{e_r} \qquad (0 < p_1 < p_2 < \cdots < p_r).$$

对每个 $P=p_i$, 有 $P^e|(q-1)$, 所以 $x^{P^e}=1$ 的全部根都是 $x^{q-1}=1$ 的根, 因此它们都在 F 中. 这个方程 $x^{P^e}=1$ 的所有 P^e 个不同的根中, 恰有 P^{e-1} 个根满足方程 $x^{P^{e-1}}=1$, 所以 F 至少包含 $x^{P^e}=1$ 的一个根 $c=c_i$, 它不满足 $x^{P^{e-1}}=1$. 于是这个元素 c_i 在 F 的乘法群中的阶数为 $p_i^{e_i}$. 乘积 $c_1c_2\cdots c_r$ 是 $q-1$ 阶元素 (参看下面的习题 8), 这正是所要求的.

定理 7　*每个特征为 p 的有限域有一个自同构 $a\longleftrightarrow a^p$.*

证明　从特征为 p 的域的一般讨论中我们知道, 对应 $a\longmapsto a^p$ 把 F 同构地映射到它元素的 p 次幂组成的集合 (13.7 节定理 13). 因为这个对应是一一的, 所以 q 个元素 a 恰好给出 q 个 p 次幂, 那么它们一定包含整个域 F. 因此 $a\longmapsto a^p$ 把 F 映射到整个 F 上.

推论　*在特征为 p 的有限域中, 每个元素有 p 次根.*

有限域的另外一些性质将在习题中叙述.

习　题

1. 证明: 对每个正整数 n, 都存在一个 $\mathbf{Z}_p$ 上 n 次不可约多项式.
2. 证明: 包含 $\mathbf{Z}_p$ 的每个有限域是 $\mathbf{Z}_p$ 的单扩张.
3. 证明: 有限域的每个有限扩张是单扩张.
4. (a) 用次数证明: $GF(p^n)$ 的任意子域有 p^m 个元素, 这里 $m|n$.
   (b) 证明: 如果 $m|n$, 那么 $(p^m-1)|(p^n-1)$.
   (c) 用 (b) 证明: 如果 $m|n$, 那么 $GF(p^n)$ 有一个含有 p^m 个元素的子域.
5. 证明: p^n 阶伽罗瓦域的所有子域组成的格与 n 的所有正因子组成的格同构.
6. 证明: 在 $GF(p^n)$ 中自同构 $a\longleftrightarrow a^p$ 的阶为 n.
7. 证明: 如果 m 与 F 的特征 p 互素, 那么在 F 上存在一个 m 次本原单位根. (*提示*: 应用证明定理 6 时用过的方法. 这个方法可以应用到特征为 ∞ 的域上吗?)
8. 证明: 在阿贝耳群中, 每个元素 c_i 的阶是不同素数的幂 $p_i^{e_i}$, 这些元素的乘积 $c_1c_2\cdots c_r$ 的阶恰好是 $p_1^{e_1}\cdots p_r^{e_r}=h$. $\left(\text{提示: 指出这个乘积的阶数可整除 } h\text{, 但对每个 } i\text{, 它不能整除 } \dfrac{h}{p_i}.\right)$
9. (a) 用基本原理证明: 模 p 的非零整数 (在 $\mathbf{Z}_p$ 中) 组成的乘法群是循环群.
   (b) 设 ξ 是有理数域 $\mathbf{Q}$ 上 p 次本原单位根, 利用 (a) 来证明: $\mathbf{Q}(\xi)$ 在 $\mathbf{Q}$ 上的伽罗瓦群是 $p-1$ 阶循环群.
10. (a) 证明: 在阶数为 $q=p^n$ 的任意有限域中, 由完全平方组成的集合 S 的基数至少是 $\dfrac{q+1}{2}$.
    (b) 验证: 对任意 $a\in S$, 集合 $S\cap(a-S)$ 不可能是空集.
    (c) 推出每个元素是两个平方之和.

15.4 伽罗瓦群

群不仅可以用来表示几何图形的对称，而且还可以表示代数系统的对称. 例如，复数域 $\mathbf{C}$ 相对于实数域而言有两种对称：一个是恒等，另一个是同构 $a+b\mathrm{i} \longleftrightarrow a-b\mathrm{i}$，这个同构把每个数映射到它的复共轭. 这种一个域到自身的同构称为自同构. 一般地，域 K 的自同构 T 是集合 K 到它自身的双射，并保持和与积，即对 K 中所有 a 和 b，有

$$(a+b)T = aT + bT, \quad (ab)T = (aT)(bT). \tag{5}$$

两个自同构 S 和 T 的乘积 ST 也是一个自同构，并且自同构的逆仍然是自同构. 因此

定理 8 *域 K 的所有自同构组成的集合在乘积之下构成一个群.*

设 K 是 F 的扩张，并考虑这样一些自同构 T，它对 F 中每个元素 a，有 $aT=a$. 也就是说这些自同构保持 F 中任何元素都不变. 在 K 的整个自同构群中，它们构成一个子群，称为 K 在 F 上的*自同构群*. 例如，$\mathbf{C}$ 在 $\mathbf{R}$ 上的自同构群由两个自同构 $a+b\mathrm{i} \longmapsto a+b\mathrm{i}$ 和 $a+b\mathrm{i} \longmapsto a-b\mathrm{i}$ 组成.

定义 *域 K 在子域 F 上的自同构群是由保持 F 的元素不变的 K 的那些自同构组成的群.*

最重要的特例是代数数域在有理数域 $\mathbf{Q}$ 上的自同构群，但是在我们考虑具体例子之前，先让我们确定代数数在自同构之下可能的像.

定理 9 *域 F 的有限扩张 K 的任意自同构 T 把 K 的每个元素 u 映射到 u 在 F 上的共轭元素 uT.*

这个定理断言，u 和它的像 uT 都满足 F 上同一个不可约方程. 为证明这一点，设给定的元素 u 在 F 上是代数的，它满足一个系数在 F 中的首一不可约多项式方程 $p(x)=x^n+b_{n-1}x^{n-1}+\cdots+b_0=0$. 根据公式 (5)，自同构 T 保持所有有理关系，并保持每个 b_i 固定，因此由 $p(u)=0$ 得到

$$(u^n+b_{n-1}u^{n-1}+\cdots+b_0)T = (uT)^n + b_{n-1}(uT)^{n-1}+\cdots+b_1(uT)+b_0 = 0.$$

这个方程表明，uT 也是 $p(x)$ 的根，因此 uT 是 u 的共轭.

例 1 考虑有理数域上的四次域[①]，$K=\mathbf{Q}(\sqrt{2},\mathrm{i})$，它是由 $\sqrt{2}$ 和 $\mathrm{i}=\sqrt{-1}$ 生成的. 整个域 K 是中间域 $F=\mathbf{Q}(\mathrm{i})$ 上的二次扩张，它是由 $x^2=2$ 的两个共轭根 $\pm\sqrt{2}$ 中的任意一个生成的. 根据 14.3 节定理 6，存在 K 的自同构 S，把 $\sqrt{2}$ 映射到 $-\sqrt{2}$，并

① 同 14.5 节中一样，我们可以把这个域看作 x^4-2x^2+9 的根域.

保持 $\mathbf{Q}(\mathrm{i})$ 中元素固定. 也就是说, 共轭根 $\sqrt{2}$ 和 $-\sqrt{2}$ 在代数上没有什么差别, S 作用到 K 的任意元素 u 上, 是

$$(a+b\sqrt{2}+c\mathrm{i}+d\sqrt{2}\mathrm{i})S=a-b\sqrt{2}+c\mathrm{i}-d\sqrt{2}\mathrm{i}, \tag{6}$$

这里我们通过基 $1,\sqrt{2},\mathrm{i},\sqrt{2}\mathrm{i}$ 把 K 的每个元素写出来 (参看 14.5 节). 通过类似的论证, 存在一个自同构 T, 它保持 $\mathbf{Q}(\sqrt{2})$ 的元素固定, 并把 i 映射到 $-\mathrm{i}$. 那么有

$$(a+b\sqrt{2}+c\mathrm{i}+d\sqrt{2}\mathrm{i})T=a+b\sqrt{2}-c\mathrm{i}-d\sqrt{2}\mathrm{i}, \tag{7}$$

所以 T 就是把每个数映射到它的复共轭, 乘积 ST 是 K 的第三个自同构. 这些自同构作用到 $\sqrt{2}$ 和 i 上的效果可以列表如下：

$$S:\begin{cases}\sqrt{2}\longmapsto-\sqrt{2}\\ \mathrm{i}\longmapsto\mathrm{i},\end{cases}\qquad T:\begin{cases}\sqrt{2}\longmapsto\sqrt{2}\\ \mathrm{i}\longmapsto-\mathrm{i}\end{cases}$$

$$ST:\begin{cases}\sqrt{2}\longmapsto-\sqrt{2}\\ \mathrm{i}\longmapsto-\mathrm{i},\end{cases}\qquad I:\begin{cases}\sqrt{2}\longmapsto\sqrt{2}\\ \mathrm{i}\longmapsto\mathrm{i}.\end{cases}$$

我们断言, I,S,T 和 ST 是 K 在 $\mathbf{Q}$ 上的仅有的自同构. 根据定理 9, 任意其他的自同构 U 一定把 $\sqrt{2}$ 映射到共轭数 $\pm\sqrt{2}$, 把 i 映射到共轭数 $\pm\mathrm{i}$, 恰好有四种可能性, 就是上面表中列出的 I,S,T 和 ST. 因此 U 作用到生成元 $\sqrt{2}$ 和 i 上的效果一定同这四种自同构之一是一致的, 因此它作用到整个域上的效果也是一致的. 于是 $U=I$, S,T 或 ST.

这些自同构的乘法表可以直接由上面列出的作用到 $\sqrt{2}$ 和 i 的表求出. 它是

$$S^2=I,\qquad T^2=I,\qquad ST=TS. \tag{8}$$

这完全像四群的乘法表 (见 6.7 节), 于是我们得出结论：$\mathbf{Q}(\sqrt{2},\mathrm{i})$ 的自同构群与四群 $\{I,S,T,ST\}$ 同构.

定义　*如果 $N=F(u_1,\cdots,u_n)$ 是多项式 $f(x)=(x-u_1)\cdots(x-u_n)$ 的根域, 那么 N 在 F 上的自同构群称为方程 $f(x)=0$ 的伽罗瓦群, 或者称为域 N 在 F 上的伽罗瓦群.*

为了明显地描述特殊伽罗瓦群的自同构 T, 我们进行如下讨论. 设 N 是 $f(x)$ 在 F 上的根域, 那么 T 把 $f(x)$ 的根映射到 $f(x)$ 的根 (定理 9), 并把不同的根映射到不同的根. 因此 T 的作用相当于对 $f(x)$ 所有不同的根 $u_1,\cdots,u_k$ 作一个置换 ϕ, 所以

$$u_1T=u_{1\phi},\cdots,u_kT=u_{k\phi},\quad k\leqslant n. \tag{9}$$

另一方面, 根域中每个元素 w 可以表示成多项式 $w = h(u_1, \cdots, u_k)$, 其系数在 F 中. 因为 T 保持这些系数固定, 所以由 T 的性质 (9) 得到

$$[h(u_1, \cdots, u_k)]T = h(u_1T, \cdots, u_kT) = h(u_{1\phi}, \cdots, u_{k\phi}).$$

这个公式表明, T 作用到 w 上的效果完全由 T 作用到根上的效果确定, 或者说, T 由置换 (9) 唯一确定. 因为两个置换的乘积是通过相继作用相应的自同构而得到, 所以全体形为 (9) 的置换构成一个群, 与自同构群同构. 置换 (9) 只包含这样一些置换: 它保持全体根之间的所有多项式恒等式不变, 所以它对应于自同构. 如此建立的结果可以概述如下:

定理 10 *设 $f(x)$ 是 F 上任意 n 次多项式, 它在根域 $N = F(u_1, \cdots, u_k)$ 中恰有 k 个不同的根 $u_1, \cdots, u_k$. 那么 $f(x)$ 的伽罗瓦群 G 的每个自同构 T 诱导出一个作用在 $f(x)$ 的全体不同根上的置换 $u_i \longleftrightarrow u_iT$, 而且 T 由这个置换完全确定.*

推论 1 *任意多项式的伽罗瓦群与它的根的置换群同构.*

推论 2 *n 次多项式的伽罗瓦群的阶可整除 $n!$.*

例 2 根据爱森斯坦定理, 方程 $x^4 - 3 = 0$ 在域 $\mathbf{Q}$ 上是不可约的, 并有四个根 $r, \mathrm{i}r, -r, -\mathrm{i}r$, 这里 $\mathrm{i} = \sqrt{-1}, r = \sqrt[4]{3}$ 是 3 的正实四次根. 根域 $N = \mathbf{Q}(r, \mathrm{i}r, -r, -\mathrm{i}r)$ 可以看作是由 r, i 生成的, 即 $N = \mathbf{Q}(r, \mathrm{i})$. 因为 r 在 $\mathbf{Q}$ 上是四次的, i 是复数, 因此 i 在实域 $\mathbf{Q}(r)$ 上是二次的, 所以整个根域 N 在 $\mathbf{Q}$ 上的次数是 8. 根据 14.5 节定理 9, 这个扩张 N 有 8 个元素 $1, r, r^2, r^3, \mathrm{i}, \mathrm{i}r, \mathrm{i}r^2, \mathrm{i}r^3$ 构成的一组基. 因为 N 中每个元素可以表示成这些基元素的线性组合, 其系数为有理数, 所以自同构 T 的作用只要知道了 rT 和 $\mathrm{i}T$, 就马上完全确定了.

可以很容易地构造出 N 的几个自同构. 因为 N 是实域 $\mathbf{Q}(r)$ 上的二次扩张, 所以它有一个把 N 中每个数映射到它的复共轭的自同构, 因此 $rT = r, \mathrm{i}T = -\mathrm{i}$. 另一方面, N 是子域 $\mathbf{Q}(\mathrm{i})$ 的四次扩张, 它由元素 r 生成. 根据 14.3 节定理 6, N 有一个自同构 S, 把 r 映射到它的共轭 $\mathrm{i}r$, 所以 $rS = \mathrm{i}r, \mathrm{i}S = \mathrm{i}$. 由此得到, S^2 是一个自同构, 它满足 $rS^2 = \mathrm{i}^2 r, \mathrm{i}S^2 = \mathrm{i}$, 而 $rS^3 = -\mathrm{i}r, \mathrm{i}S^3 = \mathrm{i}$. 把 S 和 T 进一步组合, 我们得到 N 的八个自同构, 它们作用到生成元 r 和 i 上的效果如下:

	I	S	S^2	S^3	T	TS	TS^2	TS^3
把 r 映射到	r	$\mathrm{i}r$	$-r$	$-\mathrm{i}r$	r	$\mathrm{i}r$	$-r$	$-\mathrm{i}r$
把 i 映射到	i	i	i	i	$-\mathrm{i}$	$-\mathrm{i}$	$-\mathrm{i}$	$-\mathrm{i}$

我们还可以计算 $TS^3 = ST, S^4 = T^2 = I$, 所以这八个自同构构成一个群, 它与正方形对称群 (6.4 节) 同构. 这些自同构组成整个伽罗瓦群, 这因为任意自同构一定把 i 映射到它的共轭 $\pm\mathrm{i}$ 中的一个, 把 r 映射到共轭 $\pm r$ 或 $\pm \mathrm{i}r$. 上面这个表包括这些作用的所有八种可能的组合.

群论的许多概念可以应用到这样的伽罗瓦群 G 上. 例如, G 包含着由 S 生成的子群 $H=[I,S,S^2,S^3]$ 和由 S^2 生成的较小的子群 $L=[I,S^2]$. 子群 H 的每个自同构保持 i 不变, 因而保持子域 $\mathbf{Q}(\mathrm{i})$ 中每个元素不变. 较小的子群 L 由这样的自同构组成, 这些自同构保持较大的子域 $\mathbf{Q}(\mathrm{i},r^2)$ 中每个元素不变. 在这种意义下, 下降的子群序列 $G\supset H\supset L\supset I$ 对应着上升的子域序列 $\mathbf{Q}\subset\mathbf{Q}(\mathrm{i})\subset\mathbf{Q}(\mathrm{i},\sqrt{3})\subset\mathbf{Q}(\mathrm{i},r)$. 于是, 这一上升的子域序列给出求解已知方程的方法, 它是通过逐次添加较简单的方程 $x^2=-1, y^2=3, z^2=\sqrt{3}$ 的根来实现的. 这个例子说明了伽罗瓦群的子群对于求方程的根式解所起的作用.

自然要考虑伽罗瓦群的同态. 上述群 G 的每个自同构 U 把 i 映射到 $\pm\mathrm{i}$, 因而把域 $\mathbf{Q}(\mathrm{i})$ 中每个元素映射到同一个域中的某个元素. 这就意味着, U 诱导出 $\mathbf{Q}(\mathrm{i})$ 的一个自同构 U^*, 这里 U^* 是对 $\mathbf{Q}(\mathrm{i})$ 中的元素 w 通过等式 $wU^*=wU$ 来定义的. 对应 $U\longmapsto U^*$ 是一个同态, 它把 N 的所有自同构 U 组成的群 G 映射到 $\mathbf{Q}(\mathrm{i})$ 的自同构组成的群 G^* 上. 但是 G^* 只有两个元素: 恒等映射 I^* 和对换 i 和 $-\mathrm{i}$ 的自同构. 进一步, $U^*=I^*$ 当且仅当 U 保持 $\mathbf{Q}(\mathrm{i})$ 的每个元素不变, 也就是当且仅当 U 在子群 $H=[I,S,S^2,S^3]$ 中. 因此 $U\longmapsto U^*$ 是 G 的满同态, 它的核是 H, 因此群 G^* 与商群 G/H 同构.

习 题

1. 对 $\mathbf{Q}(\mathrm{i},r)$ 的所有子域组成的系统画一个格图.
2. 通过证明 x^4-3 没有系数在 $\mathbf{Q}$ 中的线性因子和二次因子来证明: x^4-3 在 $\mathbf{Q}$ 上是不可约的.
3. 把 x^4-3 的伽罗瓦群的每个自同构表示成它的根的置换.
4. (a) 证明: x^4-3 在 $\mathbf{Q}(\mathrm{i})$ 上是不可约的.
   (b) 描述 x^4-3 在 $\mathbf{Q}(\mathrm{i})$ 上的伽罗瓦群.
5. 用基本原理证明: x^4-3 根的置换: $r\longmapsto \mathrm{i}r, \mathrm{i}r\longmapsto -\mathrm{i}r, -\mathrm{i}r\longmapsto -r, -r\longmapsto r$, 不可能对应于一个自同构.
6. 设 $F=\mathbf{Q}(\omega)$ 是由三次复单位根 ω 生成的域. 讨论 x^3-2 在 F 上的伽罗瓦群, 包括: 确定这个根域的次数, 用纯群论的语言来描述伽罗瓦群, 把每个自同构表示成置换.
7. 按照习题 6 的内容讨论 x^5-7 在 $\mathbf{Q}(\xi)$ 上的伽罗瓦群, 这里 ξ 是五次本原单位根.
8. 证明: 有限域的伽罗瓦群是循环群.
9. 证明: 如果 ξ 是 n 次本原单位根, 那么 $\mathbf{Q}(\xi)$ 的伽罗瓦群是阿贝耳群. (提示: 任意自同构有这样的形式 $\xi\longmapsto\xi^e$.)
10. (a) 证明: 如果 K 是 $\mathbf{Q}$ 的扩张, 那么 K 的每个自同构保持 $\mathbf{Q}$ 的每个元素不变.
    (b) 叙述并证明对于特征为 p 的域有类似的结果.

15.5 可分多项式与不可分多项式

由于存在所谓的不可分不可约多项式或不可分元素 —— 即这些元素是 n 次代数的, 但它的共轭元素的个数小于n—— 伽罗瓦群的一般讨论就变得复杂了. 对某些特征为 p 的域, 这种复杂化就出现了, 这可以用简单的例子加以说明.

设 $K = \mathbf{Z}_p(u)$ 表示模 p 整数的域 $\mathbf{Z}_p$ 的单超越扩张, 并设 F 表示由 $u^p = t$ 生成的 K 的子域 $\mathbf{Z}_p(u^p)$. 于是, F 是由 $\mathbf{Z}_p$ 上的超越元素 t 的所有有理形式组成. 原来的元素 u 满足 F 上的一个多项式方程 $f(x) = x^p - t = 0$. 这个多项式 $f(x)$ 在 $F = \mathbf{Z}_p(t)$ 上实际上是不可约的, 这是因为如果 f 在 $\mathbf{Z}_p(t)$ 上可约, 根据高斯引理 (3.9 节), 在 t 的多项式整环 $\mathbf{Z}_p[t]$ 上, f 是可约的, 但是, 由于 $f(x) = x^p - t$ 对于 t 来说是线性的, 所以这样的因式分解 $f(x) = g(x,t)h(x,t)$ 是不可能的. 因此 $f(x)$ 的根 u 在 F 上的次数是 p. 但是 $f(x)$ 在 K 上有因式分解

$$f(x) = x^p - u^p = (x-u)^p. \tag{10}$$

因此它只有一个根 u, 并且 u(虽然它的次数 $p > 1$) 除了它本身之外没有其他共轭元素.

我们可以用下面的术语来描述上述情况.

定义 域 F 上的一个 n 次多项式 $f(x)$, 如果它在某个根域 $N \supseteq F$ 中有 n 个不同的根, 那么称它在域 F 上是可分的; 否则, 称 $f(x)$ 是不可分的. 如果有限扩张 $K \supseteq F$ 中每个元素在 F 上都满足一个可分多项式方程, 那么称 K 在 F 上是可分的.

容易检验给定的多项式 $f(x) = a_0 + a_1x + \cdots + a_nx^n$ 是否是可分的. 这也就是, 首先根据公式 (参看 3.1 节习题 7)

$$f'(x) = a_1 + (2 \times a_2)x + \cdots + (n \times a_n)x^{n-1} \tag{11}$$

来定义 $f(x)$ 的形式导数 $f'(x)$, 这里 $n \times a_n$ 表示 a_n 的 n 次自然倍数(见 13.7 节). 如果这些系数都在实数域中, 那么这种导数与微积分学中建立的普通导数是一致的. 从形式导数的定义 (11) 出发, 不用任何极限概念, 我们可以推导出很多微分法则, 例如

$$(f+g)' = f' + g', \quad (fg)' = fg' + f'g, \quad (f^m)' = mf^{m-1}f',$$

等等.

现在在任意根域 N 上把 $f(x)$ 分解成不同线性因子的幂,

$$f(x) = c(x-u_1)^{e_1}\cdots(x-u_k)^{e_k} \quad (c \neq 0). \tag{12}$$

把 (12) 的两边形式地微分, 我们看到, $f'(x)$ 是 $ce_1(x-u_1)^{e_1-1}(x-u_2)^{e_2}\cdots(x-u_k)^{e_k}$ 与 $k-1$ 项每项都包含 $(x-u_1)^{e_1}$ 作为因子的和. 因此当 $e_1>1$ 时, $x-u_1$ 可整除 $f'(x)$, 而当 $e_1=1$ 时, $x-u_1$ 就不能整除 $f'(x)$. 对 $e_2,\cdots,e_k$ 重复上述推理, 我们得到 $f(x)$ 和 $f'(x)$ 有公因子, 除非 $e_1=e_2=\cdots=e_k=1$, 即除非 $f(x)$ 是可分的; 因此 $f(x)$ 在 N 上是可分的当且仅当 $f(x)$ 和它的形式导数 $f'(x)$ 是互素的.

$f(x)$ 和 $f'(x)$ 的最大公因式可像第 3 章中那样用 $F[x]$ 中的欧几里得算法直接计算出来; 当 F 扩张到较大的域上时, 它们的最大公因式并不改变. 于是我们得到

定理 11　*设 $f(x)$ 是域 F 上任意多项式, 用欧几里得算法计算 $f(x)$ 和它的形式导数 $f'(x)$ 的最大公因式 $d(x)$(首一多项式). 如果 $d(x)=1$, 那么 $f(x)$ 是可分的; 否则 $f(x)$ 是不可分的.*

如果 $f(x)$ 是不可约的, 那么除非 $f(x)$ 整除 $g(x)$, 总有 $\text{g.c.d.}(f(x),g(x))=1$, 并且 $f(x)$ 不能整除任意较低次的非零多项式. 因此得到

推论 1　*任意不可约多项式是可分的, 除非它的形式导数是零.*

推论 2　*特征为 ∞ 的域上的任意不可约多项式是可分的.*

这是因为, 当 $n>0, a_n\neq 0$ 时, $f'(x)=n\times a_nx^{n-1}+\cdots\neq 0$.

进一步推论是: 如果 F 的特征是 ∞, 那么任意 n 次不可约多项式 $f(x)$ 的根域恰恰包含 $f(x)$ 的 n 个不同的共轭根. 而且, 在特征为 ∞ 的域上的任意代数数满足一个不可约的因而也是可分的方程, 所以特征为 ∞ 的域的任意代数扩张, 在上述意义之下是可分扩张.

推论 2 的结果对于素特征的域是不成立的. 例如, 在本节开始所提到的不可约多项式 $f(x)=x^p-t$ 有形式导数 $(x^p-t)'=p\times x^{p-1}=0$.

习　题

1. 不用定理 11, 证明: $\mathbf{Q}$ 上二次不可约多项式的根是不同的.
2. 设 $f(x)$ 是有理系数多项式, 而 $d(x)$ 是 $f(x)$ 和 $f'(x)$ 的最大公因式. 证明: $\dfrac{f(x)}{d(x)}$ 是与 $f(x)$ 具有相同根的多项式, 但它没有重根.
3. (a) 证明: 如果 $f'(x)=0$, 那么 $f(x)$ 在任意域 F 上是不可分的.
   *(b) 证明: 如果在 $\mathbf{Z}_p$ 上 $f'(x)=0$, 那么对某个 $g(x)$, 有 $f(x)=[g(x)]^p$.
4. 证明: x^3-2u 在 $\mathbf{Z}_3(u)$ 上是不可分的. 证明: 它的根域的伽罗瓦群是由一个单位元素组成.
5. 用定理 11 证明: 当 $q=p^n$ 时, x^q-x 在 $\mathbf{Z}_p$ 上是可分的.
6. (a) 证明: 如果 $f(x)$ 是在特征为 p 的域 F 上的满足 $f'(x)=0$ 的一个多项式, 那么 $f(x)$ 可以写成形式 $a_0+a_1x^p+\cdots+a_nx^{np}$.
   (b) 证明: 如果 F 是有限域, 那么对某适当的 $g(x)$, 有 $f(x)=[g(x)]^p$.
   (c) 用 (b) 证明: 有限域上的每个不可约多项式是可分的.

15.6 伽罗瓦群的性质

可分多项式的根域和伽罗瓦群有两个极好的性质,现在我们把它们叙述成定理.

定理 12 *域 F 上可分多项式的伽罗瓦群的阶恰好等于它的根域的次数 $[N:F]$.*

在 15.4 节的第二个例子中我们已经看到, 多项式 $x^4-3=0$ 的根域就是这种情形.

定理 13 *在可分多项式的根域 $N\supset F$ 中, 在 N 的 (在 F 上的) 伽罗瓦群的每个自同构之下保持不变的元素恰恰就是 F 的元素.*

这个定理告诉我们一些关于伽罗瓦群正面的信息, 因为它断言, 对 N 中每个不在 F 中的元素 a, 在 G 中有一个自同构 T, 使得 $aT\neq a$.

为证明定理 12, 参考 15.2 节的引理 2, 它与域之间同构的可扩张性有关, 注意, 在这个引理中 (不像 15.5 节中那样), $f'(x)$ 并不表示 $f(x)$ 的导数.

引理 *如果 15.2 节引理 2 中的多项式 $f(x)$ 是可分的, 那么 S 可以按照 $m=[N:F]$ 种不同的方法扩张到 N.*

这个结果可以通过对 m 用数学归纳法来证明. 域 F 到 F' 的已知同构 S 的任意扩张 S^*, 按照 (2) 式把根 u 映射到 $p'(x)$ 的某个根 u', 因此 S 的每种可能的扩张可以通过上述一种构造而得到. 因为 $f(x)$ 是可分的, 所以它的 d 次因子 $p(x)$ 恰有 d 个不同的根 u. u' 的 d 种选择恰好给出 (2) 中 S^* 的 d 种选择. 根据归纳法假设, 每个这样的 S^* 可以按照 $\dfrac{m}{d}=[N:F(u)]$ 种不同的方法扩张到 N, 所以总共有 $d\left(\dfrac{m}{d}\right)=m$ 种扩张, 如断言所述.

如果 $f(x)=f'(x)$ 是 m 次可分的, 在 15.2 节引理 2 中我们令 $N=N'$, 上述引理断言, F 的恒等自同构 I 恰恰可以按照 m 种不同方法扩张到 N 的一个自同构. 但是这些自同构组成 N 在 F 上的伽罗瓦群, 这就证明了定理 12.

最后, 为了证明定理 13, 设 G 是可分多项式的根域 N 在域 F 上的伽罗瓦群, 而 K 是 N 中所有在 G 的每个自同构之下不变的元素组成的集合. 容易证明 K 是一个域, 并且 $K\supset F$. 因此 G 中每个自同构是 K 的恒等自同构 I 到 N 上的一个扩张. 因为 N 是 K 上的根域, 所以根据上述引理, 只存在 $[N:K]$ 个这样的扩张, 而根据定理 12, 总共有 $[N:F]$ 个自同构, 因此 $[N:K]=[N:F]$. 因为 $K\supset F$, 所以这就推出 $K=F$, 证明了定理 13.

上述关于扩张的引理还有另一个推论, 即根域在下述意义下总是 "正规的".

定义 *域 F 的有限扩张 N 称为在 F 上是正规的, 是指如果 F 上每个不可约多项式 $p(x)$ 在 N 中有一个根, 则它的所有根都在 N 中.*

换句话说, 每个在 F 上不可约在 N 中有一个根的多项式 $p(x)$ 在 N 上可以分解成线性因子.

定理 14 *F 的一个有限扩张在 F 上是正规的当且仅当它是 F 上某个多项式的根域.*

证明 如果 N 在 F 上是正规的, 那么在 N 中选取任意一个不在 F 中的元素 u, 并求出 u 所满足的不可约方程 $p(x)=0$. 根据正规性的定义, N 包含 $p(x)$ 的所有根, 因此 N 包含 $p(x)$ 的根域 M. 如果 N 中有不属于 M 的元素, 其中一个元素 v 满足不可约方程 $q(x)=0$, 因此 M 包含在较大的 $p(x)q(x)$ 的根域中, 等等. 因为 N 的次数是有限的, 所以这样逐次得到的根域中一定有一个是整个域 N.

反过来, 任意 $f(x)$ 的根域 N 是正规的. 假定有某个多项式 $p(x)$ 在 F 上是不可约的, 它有一个但不是全部根在 N 中, 设 w 是 $p(x)$ 的位于 N 中的根, 并把另一个不在 N 中的根 w' 添加到 N 上. 满足 $wT=w'$ 的对应 T 是单扩张 $F(w)$ 到 $F(w')$ 的同构. 域 N 是 $f(x)$ 在 $F(w)$ 上的根域; 另一方面, $N'=N(w')$ 是由 $f(x)$ 的根添加到 $F(w')$ 上生成的, 因此它是 $f(x)$ 在 $F(w')$ 上的根域, 所以根据 15.2 节引理 2, 对应 T 可以扩张成 N 到 N' 的同构. 因为 T 保持基域 F 的元素不变, 所以这两个同构的域 N 和 N' 在 F 上的次数一定相同. 但是我们已经假定了 $N'=N(w')$ 是 N 的真扩张, 所以 N' 在 F 上的次数大于 N 在 F 上的次数. 由此矛盾故得定理.

如果把这个定理证明的前一半应用到可分扩张上 (这个扩张中每个元素都满足一个可分方程), 那么所有用到的多项式 $p(x)$ 和 $q(x)$ 都是可分的. 这就证明了

推论 *F 的每个有限正规可分扩张是一个可分多项式的根域.*

特别是, 有理数域 $\mathbf{Q}$ 的每个有限正规扩张 N 自然就是可分的 (定理 11 的推论 2), 因此它是某个可分多项式的根域. 因此 N 在 $\mathbf{Q}$ 上的自同构群的阶恰好就是 N 在 $\mathbf{Q}$ 上的次数 $[N:\mathbf{Q}]$.

伽罗瓦群可以用来研究对称多项式的性质, 关于对称多项式已在 6.10 节中给出定义.

定理 15 *设 $N=F(u_1,\cdots,u_n)$ 是由 n 次可分多项式 $f(x)$ 的全部 n 个根 $u_1,\cdots,u_n$ 生成的域, 并设 $g(x_1,\cdots,x_n)$ 是 F 上 n 个未定元 $x_1,\cdots,x_n$ 的任意对称多项式形式. 那么 N 的元素 $w=g(u_1,\cdots,u_n)$ 在基域 F 中.*

证明 根据定理 10, N 的伽罗瓦群 G 的任意自同构 T, 它的作用相当于对 $f(x)$ 的根做置换 $u_i \longmapsto u_iT$. $g(x_1,\cdots,x_n)$ 的对称性意味着, 对于未定元的任意置换, 它都不变; 因此

$$w \longmapsto wT = g(u_1T,\cdots,u_nT) = g(u_1,\cdots,u_n) = w.$$

因为 w 在任意自同构 T 作用之下都不变, 所以根据定理 13, w 在 F 中.

推论 任意 n 个未定元的对称多项式 (F 上的) 可以表示成 n 个初等对称函数

$$\begin{aligned}\sigma_1 &= x_1 + x_2 + \cdots + x_n,\\ \sigma_2 &= x_1x_2 + x_1x_3 + \cdots + x_{n-1}x_n,\\ &\vdots\\ \sigma_n &= x_1x_2\cdots x_n\end{aligned} \tag{13}$$

的有理函数①(F 上的).

为简化公式, 我们只写出 $n = 3$ 情形的证明. F 上的初等对称函数 σ_1, σ_2 和 σ_3 生成一个域 $K = F(\sigma_1, \sigma_2, \sigma_3)$. 由原来三个未定元生成的域 $N = F(x_1, x_2, x_3)$ 是 K 的有限扩张, 事实上, N 的生成元 x_1, x_2, x_3 是三次多项式

$$f(x) = (x - x_1)(x - x_2)(x - x_3) = x^3 - \sigma_1x^2 + \sigma_2x - \sigma_3$$

的根, 其中系数原来就是 (13) 式给出的对称函数. 引进根域 N 在 K 上的伽罗瓦群 G. 根据定理 10, 每个自同构诱导出一个 x_1, x_2, x_3 的置换, 因此根据定理 15, x_1, x_2, x_3 的任意对称多项式在基域 K 中. 因为 $K = F(\sigma_1, \sigma_2, \sigma_3)$, 所以由此得出, 这样的对称多项式是 $\sigma_1, \sigma_2, \sigma_3$ 的有理函数.

习　题

1. 在定理 15 的推论的证明中, 证明: $N = K(x_1, x_2, x_3)$ 在 K 上的伽罗瓦群恰好是三个字母的对称群.
2. 把 $x_1^3 + x_2^3 + x_3^3$ 表示成初等对称多项式的有理函数. (也可参看 6.10 节习题 7 和习题 8.)
3. (a) 证明: 存在域 K 和子域 F, 使得 K 在 F 上的伽罗瓦群是 n 次对称群.
   (b) 证明: 在 (a) 中, K 可以选为实数域的一个子域. (提示: 利用 n 个代数无关的实数.)
4. 设 n 次多项式有 n 个根 $x_1, \cdots, x_n$, 它的判别式是 $D = \Pi(x_i - x_j)^2$, 这里的乘积是取遍满足 $i < j$ 的所有下标对.
   (a) 证明: 有理系数多项式的判别式是有理数.
   (b) 对于二次多项式, 把 D 明显地表示成系数的有理函数.
   *(c) 对三次多项式做同样的问题.
5. 证明: 如果 K 在 F 上是正规的, 并且 $F \subset L \subset K$, 那么 K 在 L 上是正规的.

15.7 子群与子域

如果 H 是域 N 的任意自同构集合, 那么 N 中所有在 H 的全部自同构之下保

① 参看 6.10 节定理 19, 在那里叙述了一个较强的结果.

持不变的元素 a(对 H 中每个 T, 有 $aT=a$) 构成 N 的一个子域. 特别是, 如果 N 是任意多项式在任意基域 F 上的根域, 并且 H 是 N 在 F 上的伽罗瓦群的任意子群, 则上述结论也是正确的.

定理 16 如果 H 是域 N 的任意有限自同构群, 而 K 是由所有在 H 之下不变的元素组成的子域, 那么 N 在 K 上的次数 $[N:K]$ 至多等于 H 的阶.

证明[①] 如果 H 的阶是 n, 那么只须证明, N 中任意 $n+1$ 个元素 $c_1,\cdots,c_{n+1}$ 在 K 上是线性相关的. 从 H 的 n 个元素 T 出发我们构造 $n+1$ 个未知数 y_i 的 n 个齐次方程的方程组

$$y_1(c_1T)+y_2(c_2T)+\cdots+y_{n+1}(c_{n+1}T)=0.$$

根据 2.3 节定理 10, 这样的方程组在 N 中总有一组不同于 $y_1=y_2=\cdots=y_{n+1}=0$ 的解. 现在选取最小的整数 m, 使得 n 个方程

$$y_1(c_1T)+y_2(c_2T)+\cdots+y_m(c_mT)=0,\quad T\in H \tag{14}$$

还有这样的解. 这组解 $y_1,\cdots,y_m$ 是由 N 的元素组成的, 并且除常数因子外, 它们是唯一的, 这是因为如果有两组不成比例的解, 那么通过适当的线性组合将得到含有 $m-1$ 个未知数的方程组的解. 不失一般性, 我们也可以假定 $y_1=1$.

现在把 H 中的任意自同构 S 作用到 (14) 式的左边. 因为 $TS=T'$ 跑遍 H 的所有元素, 所以得到方程组

$$(y_1S)(c_1T')+(y_2S)(c_2T')+\cdots+(y_mS)(c_mT')=0,\quad T'\in H,$$

除了方程排列的次序外, 它与 (14) 是一样的. 因此 $y_1S,\cdots,y_mS$ 也是 (14) 的解, 并且根据解的唯一性, 这组解就是 $ty_1,\cdots,ty_m$, 其中 t 是比例因子. 然而, 因为 $y_1=1,S$ 是自同构, 所以 $y_1S=1$, 于是 $t=1$. 我们得出结论: 对每个 $i=1,\cdots,m$, 和 H 中每个 S, 有 $y_iS=y_i$, 这就意味着, 系数 y_i 属于 H 之下不变的元素组成的子域 K 中. 方程 (14) 中取 $T=I$, 这就表明元素 $c_1,\cdots,c_m$ 在域 K 上是线性相关的, 这就证明了定理.

根据这个定理, 我们至少可以对可分多项式建立伽罗瓦群的子群与相应根域的子域之间的对应. 这个对应为把已知方程的根域问题化为平行的 (有限) 伽罗瓦群的子群问题提供了一个系统方法.

定理 17(伽罗瓦理论基本定理) 如果 G 是 F 上的可分多项式 $f(x)$ 的根域 N 的伽罗瓦群, 那么存在 G 的子群 H 和 N 中包含 F 的子域 K 之间的双射 $H\longleftrightarrow K$.

① 这个证明应归于阿廷 (Artin) 教授. 它包含着这样一个思想: 即认为伽罗瓦群只不过是有限自同构群, 与基域没有明显的关系.

如果 K 已给定, 则对应的子群 $H=H(K)$ 是由 G 中所有保持 K 的元素不变的自同构组成; 如果 H 已给定, 则对应的子域 $K=K(H)$ 是由 N 中所有在子群 H 的每个自同构之下保持不变的元素组成. 对于每个 K, 子群 $H(K)$ 是 N 在 K 上的伽罗瓦群, 它的阶等于 N 在 K 上的次数 $[N:K]$.

证明 对于给定的 K, 这样来描述 $H(K)$:

$$T \text{ 在 } H(K) \text{ 中} \quad \text{当且仅当 } bT=b(\text{对 } K \text{ 中所有 } b). \tag{15}$$

如果 S 和 T 具有这种性质, 则乘积 ST 也具有这种性质, 所以集合 $H(K)$ 是一个子群. 域 N 是 $f(x)$ 在 K 上的根域, N 在 K 上的每个自同构一定是 N 在 F 上保持 K 中每个元素不变的自同构, 因此它在子群 $H(K)$ 中. 所以根据定义, $H(K)$ 是 N 在 K 上的伽罗瓦群. 如果把定理 12 应用于这个伽罗瓦群, 就可证明 $H(K)$ 的阶数恰好是 N 在 K 上的次数.

两个不同的中间域 K_1 和 K_2 确定不同的子群 $H(K_1)$ 和 $H(K_2)$. 为了证明这一点, 选择任意一个在 K_1 中不在 K_2 中的元素 a, 并对 N 在 K_2 上的群 $H(K_2)$ 应用定理 13. 这就可以断言, $H(K_2)$ 包含某个 T 使得 $aT\neq a$. 因为 a 是在 K_1 中, 所以这个自同构 T 不在群 $H(K_1)$ 中, 于是 $H(K_1)\neq H(K_2)$.

我们现在知道, 对应 $K\longmapsto H(K)$ 是 N 的所有子域和 G 的某些子群之间的双射. 为了建立所有子域和所有子群之间的双射, 我们必须指出每个子群表现为 $H(K)$. 设 H 为 h 阶子群, $K=K(H)$ 像在定理 17 中那样定义:

$$b \text{ 在 } K(H) \text{ 中} \quad \text{当且仅当 } bS=b \ (\text{对 } H \text{ 中所有 } S). \tag{16}$$

根据定理 16, 有 $[N:K]\leqslant h$. 比较 (15) 与 (16) 我们看出, 对应于 $K=K(H)$ 的子群 $H(K)$ 必然包含原来给出的群 H, 而根据定理 12, $H(K)$ 的阶是 $[N:K]$. 因为 $[N:K]\leqslant h$, 所以这就意味着群 $H(K)$ 的阶不超过它的子群 H 的阶. 因此 $H(K)=H$, 如断言所述. 这就完成了定理的证明.

N 和 F 中间的所有子域 K 组成的集合, 对于子域之间的普通包含关系来说, 它是一个格. 如果 K_1 和 K_2 是两个子域, 它们的最大下界 (或在这个格中的交) 是交集 $K_1\cap K_2$, 它是由 K_1 和 K_2 的所有公共元素组成, 而它们的最小上界 (或在这个格中的并) 是 $K_1\vee K_2$, 它是由 K_1 和 K_2 的全体元素共同生成的 N 的子域. 例如, 如果 $K_1=F(v_1)$ 和 $K_2=F(v_2)$ 都是单扩张, 那么它们的并就是多重扩张 $F(v_1,v_2)$.

定理 18 *所有子域 $K_1,K_2,\cdots$ 组成的格, 通过定理 17 中所述的对应 $K\longmapsto H(K)$, 按照下述方式映射到由 G 的所有子群组成的格上:*

$$\text{由 } K_1\subset K_2 \quad \text{可推出 } H(K_1)\supset H(K_2) \tag{17}$$

$$H(K_1\vee K_2)=H(K_1)\cap H(K_2), \tag{18}$$

$$H(K_1 \cap K_2) = H(K_1) \vee H(K_2). \tag{19}$$

特别是，仅由单位元素组成的子群对应着整个正规域 N.

这些结果表明，这个对应把包含关系颠倒过来，把任意交映射到并，并且把并映射到交，具有这些性质的两个格之间的任意双射称为*对偶同构*.

为了证明这个定理，我们首先注意，对应于域 K 的群的定义 (15) 表明，对应于较大子域的群一定使更多的元素保持不变，因此这个群就较小，这就得到 (17). 交和并纯粹按照包含关系来定义 (见 11.7 节)，因此根据对偶原理，使包含关系颠倒的双射一定把交与并对换，这就是 (18) 和 (19) 式所断言的.

我们省略了下述进一步结果的证明.

定理 19 *满足 $N \supset K \supset F$ 的域 K 是 F 上的正规域当且仅当它所对应的群 $H(K)$ 是 N 的伽罗瓦群 G 的正规子群. 如果 K 是正规的，那么 K 在 F 上的伽罗瓦群 K 与商群 $G/H(K)$ 同构.*

这个定理的结论已经在 15.4 节的末尾解释过了，那里所举的例子是这个定理的特殊情形.

习 题

1. (a) 证明：如果 H 是域 N 的任意自同构集合，那么 N 中所有在 H 的全部自同构之下保持不变的元素构成 N 的一个子域 K.
   (b) 证明：N 在这个子域 K 上是正规的.
2. 对 $\mathbf{Q}$ 上的域 $\mathbf{Q}(\sqrt{2}, \mathrm{i})$，完整地列出它的子域与子群的对应.
3. 对 15.4 节中所讨论的 $x^4 - 3$ 的根域，做与习题 2 相同的问题.
4. 证明：$H(K)$ 在 G 中的指数等于 K 在 F 上的次数.
5. 证明：如果 N 是 F 上可分多项式 $f(x)$ 的根域，那么 N 和 F 之间的中间域个数是有限的.
6. 证明：N 和 F 之间的所有中间域 K 构成格.
7. 证明：如果 K 是特征为 ∞ 的域 F 的有限扩张，那么 K 和 F 之间的中间域个数是有限的.

*8. 证明定理 19.

*9. 在定理 17 意义下的两个子域 K_1 和 K_2，如果存在 N 在 F 上的一个自同构 T 把 K_1 映射到 K_2，则称 K_1 和 K_2 是*共轭的*. 证明：K_1 和 K_2 是共轭的当且仅当 $T^{-1}H(K_1)T = H(K_2)$(也就是说，当且仅当 $H(K_1)$ 和 $H(K_2)$ 是 G 的共轭子群).

15.8 三次不可约方程

伽罗瓦理论可以用来证明，关于用根式解方程的各种经典问题的不可解性. 作

为这个方法的简单例子, 我们来考虑有名的具有实根的三次不可约方程.

一个三次方程可以取为如下形式 (见 5.5 节 (17) 式)

$$f(y)=y^3+py+q=(y-y_1)(y-y_2)(y-y_3), \tag{20}$$

该方程含有实系数 p 和 q, 并有三个实的或复的根 y_1,y_2,y_3. 系数 p 和 q 可以表示成这些根的对称函数, 因为当把 (20) 乘出来时, 我们便得到

$$0=y_1+y_2+y_3,\quad p=y_1y_2+y_1y_3+y_2y_3,\quad q=-y_1y_2y_3. \tag{21}$$

引进三次方程的判别式是很重要的, 它用下面的公式来定义:

$$D=[(y_1-y_2)(y_1-y_3)(y_2-y_3)]^2. \tag{22}$$

任意两个根的置换不改变 D, 所以 D 是 y_1,y_2 和 y_3 的对称多项式. 根据定理 15, 可推出 D 可表示成域 $F=\mathbf{Q}(p,q)$ 中的元素, F 是由系数 p 和 q 生成的. 这个表达式像 5.5 节 (24) 中的一样, 是

$$D=-4p^3-27q^2. \tag{23}$$

这个等式是 y_1,y_2,y_3 的多项式恒等式, 并可用方程 (21) 和 (22) 直接验证.

定理 20 *具有正判别式的实三次方程有三个实根; 如果 $D=0$, 则至少有两个根是相等的; 如果 $D<0$, 则有两个根是虚根.*

只要考察各种类型的根对 D 的公式 (22) 有什么影响, 就可以验证我们的定理. 如果所有的根都是实的, D 显然是正的; 而如果两个根相等, 则 $D=0$. 最后, 假定一个根 $y_1=a+bi$ 是虚数 ($b\neq 0$), 那么它的复共轭 $y_2=a-bi$ 也一定是一个根 (5.4 节), 而第三个根是实根. 在 (22) 中, $y_1-y_2=(a+bi)-(a-bi)=2bi$ 是纯虚数, 而

$$(y_1-y_3)(y_2-y_3)=(y_1-y_3)(y_1^*-y_3)=(y_1-y_3)(y_1-y_3)^*$$

是一个实数, 所以判别式 D 就是负的. 这恰好给出定理中所列出的几种可能情形.

定理 21 *如果三次多项式 (20) 在 $F=\mathbf{Q}(p,q)$ 上是不可约的, 有根 y_1,y_2,y_3 和判别式 D, 那么它的根域 $F(y_1,y_2,y_3)$ 是 $F(\sqrt{D},y_1)$.*

证明 根据 D 的定义 (22), 这个根域一定包含 $\sqrt{D}$, 因此只剩下证明根 y_2,y_3 包含在域 $K=F(\sqrt{D},y_1)$ 中. 在域 K 中, 这个三次多项式有一个线性因子 $y-y_1$, 所以剩下来的二次因子

$$(y-y_2)(y-y_3)=y^2-(y_2+y_3)y+y_2y_3, \tag{24}$$

它的系数还在 K 中. 把 y_1 代入 (24) 中, 则 $(y_1-y_2)(y_1-y_3)$ 在 K 中, 所以

$$y_2-y_3=\pm\frac{\sqrt{D}}{(y_1-y_2)(y_1-y_3)}$$

在 K 中. 但是 (24) 的系数 y_2+y_3 也在 K 中. 因为 y_2+y_3 和 y_2-y_3 都在 K 中, 所以 y_2 和 y_3 也在 K 中. 这就证明了定理.

现在考虑在其系数域上是不可约的三次多项式, 它有三个实根. 5.5 节的公式 (19) 给出这些根是 $y=z-\frac{p}{3z}$, 其中

$$z^3=-\frac{q}{2}+\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}=-\frac{q}{2}+\sqrt{-\frac{D}{108}}.$$

(这里我们用了 D 的表达式 (23).) 因为这些根是实的, 所以 D 是正的 (定理 20), 因此上面公式中的平方根是虚数. 于是, 这个公式是通过复数给出实根 y!

多少年来, 这被认为是这组公式的一个严重缺点, 并有很多数学家尽力寻找三次方程实根的其他只包含实根式 (平方根、立方根或高次方根) 的公式. 但是这种探求都落空了, 这是由于下面定理的缘故.

定理 22 *如果一个三次多项式有实根, 并在由它的系数生成的域 $F=\mathbf{Q}(p,q)$ 上是不可约的, 那么不存在求三次多项式实根的有理公式, 这个公式是通过 F 上的实根式来表示的.*

在证明这个定理之前, 我们更普遍地讨论根式 $\sqrt[m]{a}=a^{\frac{1}{m}}$ 的性质. 如果 m 是复合数, 满足 $m=rs$, 那么 $a^{\frac{1}{m}}=(a^{\frac{1}{r}})^{\frac{1}{s}}$, 等等, 所以任意根式可以通过一系列素指数的根式而得到. 在后面的情形中, 我们可以确定通过添加一个根式所得到的域的次数.

引理 *实域*[①]*K 上的 r(素数) 次多项式 x^r-a 或者在 K 上不可约, 或者在 K 中有根.*

证明 把 r 次本原单位根 ξ 添加到 K 上, 然后再把 x^r-a 的一个根 u 添加上去. 所得到的扩张 $K(\xi,u)$ 包含多项式 x^r-a 的 r 个根 $u,\xi u,\xi^2u,\cdots,\xi^{r-1}u$, 因此它是这个多项式的根域, 这个多项式有因子分解式

$$x^r-a=(x-u)(x-\xi u)(x-\xi^2u)\cdots(x-\xi^{r-1}u).$$

假设 x^r-a 在 K 上有正次数 $m<r$ 的真因子 $g(x)$. 那么这个因子 $g(x)$ 是 x^r-a 在 $K(\xi,u)$ 上的 m 个线性因子之积, 所以 $g(x)$ 中的常数项 b 是 m 个根 ξ^iu 的乘积. 因此 $b=\xi^ku^m$, 对某个整数 k, 并且有

$$b^r=(\xi^ku^m)^r=(\xi^r)^k(u^r)^m=(u^r)^m=a^m.$$

由此我们可以在 K 中求出 a 的 r 次根, 这是因为 $m<r$ 与 r 互素, 于是存在整数 s 和 t 使得 $sm+tr=1$(1.7 节 (13)), 所以

① 实域是指其元素为实数的任意域. 这个引理对于任意域都是正确的, 当 K 的特征是 r 时证明必须稍微修正一下.

$$b^{sr} = a^{sm} = a^{1-tr} = \frac{a}{a^{tr}},$$

则有 $a = (b^s a^t)^r$. 于是由 $x^r - a$ 在 K 上是可约的这一假定得出 $x^r - a$ 在 K 中有一个根 $b^s a^t$. 证毕

我们现在可以证明定理 22, 为此, 假定这个结论是错误的. 那么这个三次多项式的某个根可以通过实根式表示, 这就是说, 根 y_1 在某个域 $L = F(\sqrt[r]{a}, \sqrt[s]{b}, \cdots)$ 中, 域 L 是在 F 上添加实根式而生成的. 因为 D 是正的, 把实根式 $\sqrt{D}$ 添加到这个域上得到另一个实域 $K = L(\sqrt{D})$. 根据定理 21, 这三次多项式的全部根都在这个域中, 所以它们都可以用含有实根式的公式表示. 域 K 可以由有限多个根式得到. 如果首先添加 $\sqrt{D}$, 这就相当于说, K 是下面域的有限链中最后一个

$$F \subset K_1 \subset K_2 \subset K_3 \subset \cdots \subset K_n = K, \tag{25}$$

其中

$$K_1 = F(\sqrt{D}), \quad K_{i+1} = K_i(a_i^{\frac{1}{r_i}}), \quad i = 1, \cdots, n-1, \tag{26}$$

这里每个 a_i 在 K_i 中, 每个 r_i 是素数. 去掉额外的域, 我们可以假定实根 $a_i^{\frac{1}{r_i}}$ 不在域 K_i 中, 根据引理, 这就意味着 $x^{r_i} - a_i$ 在 K_i 上是不可约的, 因此 K_{i+1} 的次数是 $[K_{i+1} : K_i] = r_i$.

根据假定, 三次多项式的根在 K 中, 它们并不在 F 中或者 $F(\sqrt{D})$ 中, 这因为三次多项式在 F 上是不可约的. 那么在链 (25) 中存在第一个包含三次多项式的一个根 (比如说 y_1) 的域 K_{j+1}. 在前一个域 K_j 上, 已知的三次多项式一定是不可约的, 如果不然, 它在 K_j 上将有线性因子 $y - y_i$, 这与 K_j 不包含任何一个 y_i 这一事实相矛盾. 那么扩张

$$K_{j+1} = K_j(a^{\frac{1}{r}}), \quad a = a_j, \quad r = r_j \tag{27}$$

的次数是 r, 并且包含一个元素 y_1, y_1 在 K_j 上的次数是 3. 根据 14.5 节定理 9 的推论 2, 有 $3|r$, 所以素数 r 一定是 3, 于是我们在 (27) 中讨论三次根 $\sqrt[3]{a}$. 这个域 K_{j+1} 是在 K_j 上添加 y_1 而生成的, 它包含 $\sqrt{D}$, 因此根据定理 21, 它包含这个三次多项式的全部根. 所以 K_{j+1} 是给定的三次多项式在 K_j 上的根域. 根据定理 14, 它作为一个根域, 是正规的. 因为它包含 K_j 上不可约多项式 $x^3 - a$ 的一个根 $a^{\frac{1}{3}}$, 所以它一定包含这个多项式的全部根. 其他两个根是 $\omega a^{\frac{1}{3}}$ 和 $\omega^2 a^{\frac{1}{3}}$, 所以 K_{j+1} 也包含三次复单位根 ω. 这与 $K_{j+1} \subset K$ 是实域的假定相违背. 定理证完.

习 题

1. 验证判别式 (23).
2. 依据定理 21 的方法, 用 y_1 和 $\sqrt{D}$ 明显地表示出三次多项式的所有根.

*3. 关于三次方程的讨论, 有哪些可以应用到 $\mathbf{Z}_3$ 上的三次方程?

4. 证明：一个多项式 x^n-a, 如果在特征为 ∞ 的域 F 上有次数与 n 互素的因子, 那么它在 F 中有根.

5. 证明：如果 F 是特征为 ∞ 的域, 它包含所有 n 次单位根, 那么次数 $[F(a^{\frac{1}{n}}):F]$ 是 n 的因子.

6. 考虑三次不可约多项式 (20) 在 $F=\mathbf{Q}(p,q)$ 上的伽罗瓦群 G. 证明：如果 D 是 F 的一个数的平方, 那么 G 是三个字母的交错群, 否则它是对称群.

15.9 五次方程的不可解性

在本节中, F 表示复数域中包含所有单位根的一个子域, K 表示 F 的各种有限扩张.

假设 $K=F(a^{\frac{1}{r}})$ 是由 F 和 $a\in F$ 的单个 r 次根 $a^{\frac{1}{r}}$ 生成的, 这里 r 是素数. 像在第 5 章那样, $x^r=a$ 的其他根是 $\xi a^{\frac{1}{r}},\cdots,\xi^{r-1}a^{\frac{1}{r}}$, 这里 ξ 是 r 次本原单位根, 因此它也在 F 中. 根据 15.8 节的引理, 除非 $K=F$, 多项式 x^r-a 在 F 上是不可约的, 所以有一个 K 的自同构 S 把根 $a^{\frac{1}{r}}$ 映射到根 $\xi a^{\frac{1}{r}}$. 这个自同构的幂 $I,S,S^2,\cdots,S^{r-1}$ 把 $a^{\frac{1}{r}}$ 分别映射到方程 $x^r=a$ 的每一个根, 因此这些幂包含 K 在 F 上的全部自同构. 于是我们得出结论：*K 在 F 上的伽罗瓦群是循环群.*

更一般地, 假设 K 在 F 上是正规的, 并可由 F 通过一系列单扩张得到, 每个单扩张只是在前一个 F 的扩张上添加一个 n_i 次根得到的. 这就意味着, 存在一系列中间域 K_i

$$F=K_0\subset K_1\subset K_2\subset\cdots\subset K_s=K \tag{28}$$

满足 $K_i=K_{i-1}(x_i)$, 这里 $x_i^{n_i}\in K_{i-1}$. 不失一般性, 我们可以假定每个 n_i 是素数. 我们把这样的 K 称为 F 的*根式扩张*. 因为 K 是正规的, 它是多项式 $f(x)$ 在 F 上的根域, 所以它也是同一个多项式 $f(x)$ 在 K_1 上的根域 —— 所以 (根据定理 14) 它在 K_1 上也是正规的. 但是根据上一节, K_1 在 F 上是正规的. 所以 K 在 F 上的每个自同构诱导出 K_1 在 F 上的一个自同构, 并且自同构的乘法也相同. 进一步, 根据 15.2 节引理 2, K_1 在 F 上的每个自同构可以扩张成 K 在 F 上的一个自同构. 因此对应是从 K 在 F 上的伽罗瓦群到 K_1 在 F 上的伽罗瓦群的*满同态*, 这很像 15.4 节末尾所描述的那样. 此外, 在这个满同态之下, 诱导出 K_1 在 F 上的恒等自同构的元素刚好是由 K 在 K_1 上的自同构定义的. 这就表明, 在这个满同态之下, K 在 F 上的伽罗瓦群 $G(K/F)$ 被映射到 $G(K_1/F)$ 上. 因此 $G(K_1/F)$ 与商群 $G(K/F)/G(K/K_1)$ 同构. 把这个结果同上一节的结果结合起来, 我们推出 $G(K/K_1)$ 是 $G(K/F)$ 的正规子群, 并且商群 $G(K_1/F)$ 是循环商群.

现在对 s 用归纳法. 根据定义, K 是 K_1 的根式扩张; 如上所述, 它在 K_1 上也

是正规的. 因此前面的推理可以再应用到 $G(K/K_1)$ 上, 证明 $G(K/K_2)$ 是 $G(K/K_1)$ 的正规子群, 它们构成循环商群 $G(K_2/K_1)$. 把这个论证重复 s 次, 并用 S_i 表示子群 $G(K/K_i)$, 我们就得出下面基本结果.

定理 23 *设 K 是 F 的任意正规根式扩张, 那么 K 在 F 上的伽罗瓦群 G 包含一个子群序列 $S_0 = G \supset S_1 \supset S_2 \supset \cdots \supset S_s$, 其中每个子群是它前一子群的正规子群, 并且商群 S_{i-1}/S_i 是循环商群, 而 S_s 仅由 I 组成.*

这表明 K 在 F 上的伽罗瓦群 G 在下面定义之下是可解的.

定义 *有限群 G 是可解的当且仅当它包含一个子群链 $S_0 = G \supset S_1 \supset S_2 \supset \cdots \supset S_s = I$ 使得对所有的 k, 有 (i)S_k 是 S_{k-1} 中的正规子群; (ii) S_{k-1}/S_k 是循环群.*

关于抽象可解群, 有很大一部分是知道的, 例如, 任意一个群, 如果它的阶可被少于三个的不同素数整除, 那么它是可解群(波恩赛德 (Burnside)). 甚至知道, 每个奇数阶群是可解的 (费特–汤姆森 (Feit-Thompson)). 然而, 我们只满足于下面的简单事实.

引理 1 *有限可解群 G 的任意满同态像 G' 本身是可解的.*

证明 设 G 有一个像可解性定义中所描述的那样的子群链 $S_0 = G \supset S_1 \supset S_2 \supset \cdots \supset S_s = I$, 并设 $S'_0 = G', S'_1, \cdots, S'_s = I'$ 是它们的满同态像. 那么, 每个 S'_k, 如果它包含 x' 和 y', 则必包含 $x'y' = (xy)'$ 和 $(x')^{-1} = (x^{-1})'$(x, y 分别是 x' 和 y' 在 S_k 中的像源), 所以 S'_k 是 G' 的子群. 进一步, 如果 a 在 S_{k-1} 中, x 在 S_k 中, 那么 S_k 在 S_{k-1} 中的正规性意味着 $a^{-1}xa$ 在 S_k 中, 因此 $(a')^{-1}x'a' = (a^{-1}xa)'$ 在 S'_k 中. 因为 a' 可以是 S'_{k-1} 的任意元素, 所以这就证明了 S'_k 在 S'_{k-1} 中是正规的. 最后, 因为 S_{k-1} 是由 S_k 的某单个陪集的各次幂 $(S_ka)^n = S_ka^n$ 组成 (S_{k-1}/S_k 是循环的), 所以 S'_{k-1} 是这个陪集的像的各次幂 $S'_k(a')^n = (S'_ka')^n$ 组成, 于是 S'_{k-1}/S'_k 是循环的. 所以这些子群组成的链 $S'_0 \supset S'_1 \supset S'_2 \supset \cdots \supset S'_s = I'$ 具有使 G' 是可解群的全部性质, 正如引理 1 所要求的. 证毕

现在让我们来定义, 系数在 F 中的一个方程 $f(x) = 0$, 如果它的根在 F 的扩张 K 中, 这个 K 可以通过在 F 上逐次添加 n 次根而得到, 我们就称 $f(x) = 0$ 在 F 上是*根式可解*的. 根据 5.5 节, 所有二次方程、三次方程和四次方程, 情况都是如此. 应看到, K 并不要求是正规的, 而只要求包含 $f(x)$ 在 F 上的根域 N. 然而, 因为可用根式表示的元素的任意共轭元素本身可用共轭根式表示, 所以 $f(x)$ 的根域 N 也一定包含在有限扩张 $K^* \supset K$ 中, K^* 在 F 上是正规的, 并是 F 的根式扩张. 这个 K^* 包含 N 作为 F 上的正规子域. 因此 K^* 在 F 上的每个自同构 S 诱导出 N 在 F 上的一个自同构 S_1, 而且对应 $S \longmapsto S_1$ 是满同态. 这就是说, K^* 在 F 上的伽罗瓦群与 N 在 F 上的伽罗瓦群是满同态的, 但是, 前者是可解的 (根据定理 23), 因此根据引理 1, 后者也是可解的, 这就证明了

定理 24 *如果系数在 F 中的一个方程 $f(x) = 0$ 是根式可解的, 那么它在 F 上的*

伽罗瓦群是可解的.

为了证明五次方程不总是根式可解的, 我们只需要找出一个方程, 它的伽罗瓦群不是可解的. 我们就来做这件事: 首先我们证明五次对称群不是可解的, 然后再列出一个五次方程, 它的伽罗瓦群是五次对称群.

定理 25　*n 个字母的对称群, 除 $n \leqslant 4$ 之外, 不是可解的.*

证明　设 $G = S_0 \supset S_1 \supset S_2 \supset \cdots \supset S_s$ 是任意子群链, 其中每个子群在前一个子群中是正规的, 并且商群 S_{k-1}/S_k 是循环商群, 我们通过对 s 用归纳法来证明, S_s 一定包含每个三字母循环 (ijk). 由此推出 $S_s > I$, 所以 G 不能是可解的.

因为 $S_0 = G$ 包含每个三字母循环, 所以我们只须用归纳法证明, 如果 S_{s-1} 包含每个三字母循环, 那么 S_s 也包含每个三字母循环. 首先注意, 如果置换 ϕ 和 ψ 都在 S_{s-1} 中, 那么它们的所谓 "换位子" $\gamma = \phi^{-1}\psi^{-1}\phi\psi$ 在 S_s 中. 为了看出这一点, 考虑 S_{s-1}/S_s 中的像 ϕ', ψ' 和 γ'. 这个商群是循环的, 也是交换群, 因此在 S_{s-1}/S_s 中有

$$\gamma' = (\phi')^{-1}(\psi')^{-1}\phi'\psi' = I',$$

这意味着 $\gamma \in S_s$. 但是在 $\phi = (ilj)$ 和 $\psi = (jkm)$ 这一特殊情形中, 这里 i, j, k 是给定的, l, m 是任意两个其他字母 (除 $n \leqslant 4$ 之外, 这样的字母是存在的), 我们有

$$\gamma = (jli)(mkj)(ilj)(jkm) = (ijk) \in S_s, \text{ 对所有的 } i, j, k.$$

这就证明了 S_s 包含每个三字母循环, 这就是所要求的.

附带说一下, 证明这个定理的更明显的形式是可能的. 我们知道, 交错群 A_n 是对称群 G 的正规子群, 所以存在一个始于 $G \supset A_n$ 的群链. 然后我们可以证明交错群 A_n(当 $n > 4$ 时), 除它本身和单位元素之外没有任何正规子群.

引理 2　*存在一个 (实) 五次方程, 它的伽罗瓦群是五个字母的对称群.*

证明　设 A 是所有代数数构成的域, 它是可数的, 并且包含所有单位根. 因此, 我们可以像 14.6 节那样相继选取五个在 A 上代数无关的实数 $x_1, \cdots, x_5$. 构成超越扩张 $A(x_1, \cdots, x_5)$. 现在设 $\sigma_1, \cdots, \sigma_5$ 是 $x_1, \cdots, x_5$ 的初等对称多项式, 并设 $F = A(\sigma_1, \cdots, \sigma_5)$. 像在定理 15 中那样, 多项式

$$f(t) = t^5 - \sigma_1 t^4 + \sigma_2 t^3 - \sigma_3 t^2 + \sigma_4 t - \sigma_5 = 0 \tag{29}$$

在 F 上的伽罗瓦群是五个字母 $x_1, \cdots, x_5$ 的对称群.

由引理 2 和定理 25 推出, 在含有一切单位根的域上存在一个 (实) 五次方程, 它的伽罗瓦群不是可解的. 现在应用定理 24, 我们得到我们的最后结论.

定理 26　*存在 (实) 五次方程, 不能用根式求解.*

习　　题

1. 证明：三个字母的对称群是可解的.
2. 证明：任意有限交换群是可解的.（提示：证明它包含一个素指数的（正规）子群.）
3. 证明：如果有限群 G 包含一个正规子群 N，使得 N 和 G/N 都是可解的，那么 G 是可解的.
4. (a) 证明：在四个字母的对称群中，三字母循环的全体换位子构成四阶正规子群.

   (b) 利用 (a) 和交错子群来证明：四个字母的对称群是可解的.
5. 证明：任意有限抽象群 G 是某适当方程的伽罗瓦群.（提示：根据凯莱定理，G 与对称群的子群同构.）

*6. (a) 证明：$x^n = a$ 的伽罗瓦群甚至在不包含单位根的域上也是可解的.

   (b) 证明：定理 24 对于任意域 F，不管它是否包含单位根，都成立.

7. 明显地证明：如果 K 是 F 的根式扩张，那么存在 K 的一个扩张 K^*，它在 F 上是正规的，并且它也是 F 的根式扩张.（上面定理 24 的证明中用到了这个事实.）
8. 证明：如果 F 包含 n 次单位根，$K = F(a^{\frac{1}{n}})$，这里 a 在 K 中，那么，甚至当 n 不是素数时，K 在 F 上的伽罗瓦群也是循环的.
9. 证明：如果 $\mathbf{Q}$ 是有理数域，f 是 (29) 式表示的特殊多项式，那么 f 在域 $\mathbf{Q}(\sigma_1, \cdots, \sigma_5)$ 上的伽罗瓦群仍然是五个字母的对称群.
10. 证明：如果 $n > 4$，那么存在一个实 n 次方程不是根式可解的.

参考文献

一般参考文献

Albert, A. A. (ed.). *Studies in Modern Algebra*(MAA Studies in Mathematics, II). Englewood Cliffs, N. J.: Prentice-Hall, 1963.

Artin, E. *Geometric Algebra.* New York: Wiley Interscience, 1957.

Birkhoff, G., and T. C. Bartee. *Modern Applied Algebra.* New York: McGraw-Hill, 1970.

Godement, Roger. *Gours d'algèbre.* Paris: Hermann, 1963.

Herstein, I. N. *Topics in Algebra.* New York: Wiley, 1964.

Jacobson, N. *Basic Algebra. I. Basic Algebra. II.* San Francisco: Freeman, 1974, 1976.

Mac Lane, Saunders, and Garrett Birkhoff. *Algebra.* New York, Chelsea, 1988.

Schreier, O., and E. Sperner. *Introduction to Modern Algebra and Matrix Theory*(English translation). New York: Chelsea, 1952.

Uspensky, J. V. *Theory of Equations.* New York: McGraw-Hill, 1948.

van der Waerden, B. L. *Modern Algebra,* I, 4th ed., and II, 5th ed. (English translation). New York: Ungar, 1966 and 1967.

数论

Hardy, G. H., and E. M. Wright. *An Introduction to the Theory of Numbers,* 4th ed. Oxford: Clarendon, 1954.

LeVeque, W. J. *Topics in Number Theory.* 2 vols. Reading, Mass.: Addison-Wesley, 1956.

Niven, lvan, and H. S. Zuckerman. *An Introduction to the Theory of Numbers.* New York: Wiley, 1960.

Rademacher, H. *Lectures on Elementary Number Theory.* New York: Wiley, 1964.

代数数论

Lang, S. *Algebraic Numbers.* Reading, Mass: Addison-Wesley, 1964.

Ribenboim, P. *Algebraic Numbers.* New York: Wiley, 1972.

Weiss, E. *Algebraic Number Theory.* New York: McGraw-Hill, 1963.

群论

Curtis, C. W., and I. Reiner. *Representation Theory of Finite Groups and Associative Algebras.* New York: Wiley Interscience, 1962.

Fuchs. L. *Abelian Groups.* Budapest: Hungarian Academy of Sciences, 1958.

Gorenstein, D. *Finite Groups.* New York: Harper & Row, 1968.

Hall, M. *The Theory of Groups.* New York: Macmillan, 1959.

Rotman, J. J. *The Theory of Groups.* Boston: Allyn & Bacon, 1965.

矩阵论

Faddaeva, V. N. *Computational Methods of Linear Algebra.* Translated by C. D. Benster. New York: Dover, 1959.

Varga, R. S. *Matrix Iterative Analysis.* Englewood Cliffs, N. J.: Prentice-Hall, 1962.

伽罗瓦理论

Artin, E. *Galois Theory,* 2nd ed. (Notre Dame Mathematical Lecture No.2). Notre Dame, Ind.: University of Notre Dame Press, 1944.

线性代数与环

Jacobson, N. *Lie Algebras.* New York: Wiley, 1962.

Jacobson, N. *The Structure of Rings,* 2nd ed. New York: American Mathematical Society, 1964.

McCoy, N. H. *The Theory of Rings.* New York: Macmillan, 1964.

代数几何

Fulton, W. *Algebraic Curves.* New York: Benjamin, 1969.

Jenner, W. E. *Rudiments of Algebraic Geometry.* New York: Oxford University Press, 1963.

Lang, S. *Introduction to Algebraic Geometry.* New York: Interscience, 1958.

Zariski, O., and P. Samuel. *Commutative Algebras.* 2 vols. New York: Van Nostrand, 1958, 1960.

逻辑学

Kleene, S. C. *Mathematical Logic.* New York: Wiley, 1967.

Mendelson, E. *Introduction to Mathematical Logic.* New York: Van Nostrand, 1964.

格论

Abbott, J. C. *Sets, Lattices and Boolean Algebras.* Boston: Allyn & Bacon, 1969.

Birkhoff, Garrett. *Lattice Theory,* 3rd ed. Providence: American Mathematical Society, 1966.

同调代数

Freyd, P. *Abelian Categories.* New York: Harper & Row, 1964.

Jans, J. P. *Rings and Homology.* New York: Holt, 1964.

Mac Lane, Saunders, *Homology.* Berlin: Springer. 1963

Mac Lane, Saunders. *Categories for the Working Mathematician.* Berlin: Springer, 1971.

泛代数

Cohn, P. M. *Universal Algebra.* New York: Harper & Row, 1965.

Grätzer, G. *Universal Algebra.* New York: Van Nostrand, 1968.

Jonsson, Bjarni. *Topics in Universal Algebra* (Lecture Notes in Mathematics No.250). Berlin: Springer, 1972.

数学符号表

$\boldsymbol{A}$	矩阵 ($\boldsymbol{B}, \boldsymbol{C}$ 等也是)
$\boldsymbol{A}^{\mathrm{T}}$	转置矩阵
$\boldsymbol{A}^*$	矩阵的复共轭
$\mathfrak{A}$	线性代数
$A_n(F)$	F 上仿射群
B	布尔代数
$\mathbf{c}$	$\mathbf{R}$ 的基数
$\mathbf{C}$	复数域
D	整环
$D[x]$	系数在 D 中的 x 的多项式形式
$D\langle x\rangle$	系数在 D 中的 x 的多项式函数
$\mathbf{d}$	正整数的集合的基数 $o(\mathbf{Z}^+)$
E_n	n 维欧几里得空间
$\boldsymbol{E}_{ij}$	特殊矩阵, (i, j) 位置的元素为 1, 其他位置为 0
$e, 1$	群的单位元素
F	域
F^n	F 上 n- 数组组成的空间
$F[x]$	系数在 F 中的 x 的多项式形式
$F(x)$	系数在 F 中的 x 的有理形式
G	群
g.1.b.	最大下界
i	$\sqrt{-1}$; 四元数单位
$\boldsymbol{I}$	恒等变换或单位矩阵; 格的最大元素
j, k	四元数单位
J	环中的理想 (H, L 等也是)
K	域
$[K:F]$	K 在 F 上的次数
$L_n(F)$	F 上全线性群
l. u. b.	最小上界
$M_n(F)$	F 上全阵代数
$\boldsymbol{O}$	零矩阵; 格的最小元素
$O_n(F)$	正交群

$o(S)$	集合 S 的基数
P	素理想; 非奇异矩阵
p, q	正素数
$Q(D)$	整环 D 的商域
$\mathbf{Q}$	有理数域
R	环
$\mathbf{R}$	实数域
S	集合; 子群; 子空间
S'	集合 S 的补
$S^{\perp}$	子空间的正交补
T	线性变换
T_A	用矩阵 $\boldsymbol{A}$ 给出的线性变换
$[u:F]$	u 在 F 上的次数
V, W	向量空间
V^*	对偶向量空间
$\boldsymbol{X}$	向量或行矩阵
z^*	共轭复数
$\mathbf{Z}$	整数环或整数群
$\mathbf{Z}_n$	模 n 整数环
$\mathbf{Z}^+$	正整数集合
$\boldsymbol{\alpha}, \boldsymbol{\beta}$	向量
$(\boldsymbol{\alpha}, \boldsymbol{\beta})$	向量内积 (点积)
$\boldsymbol{\alpha} \times \boldsymbol{\beta}$	向量外积 (向量积)
δ_{ij}	克罗内克尔符号
$\boldsymbol{\varepsilon}_i$	单位向量
ϕ, ψ	变换; 映射; 函数
$\prod$	乘积
$\sum$	求和
$\boldsymbol{\xi}, \boldsymbol{\eta}$	向量
$\mathbf{0}$	零向量
$\varnothing$	空集
$\cap, \cup$	交, 并 (集合的)
$\wedge, \vee$	交, 并 (布尔代数, 格中的)
$\in$	属于; 是 $\cdots$ 的元素
$\subset$	包含于; 是 $\cdots$ 的子集合

$<$	小于; 真包含在 …… 中
$\leqslant$	不等号
$\perp$	正交于; 垂直于
$\otimes$	张量积
$\times$	直积
$\oplus$	直和
$\circ$	二元运算
$\longmapsto$	元素的映入
$\longrightarrow$	集合的映入
∞	无限; 无穷大
$\sim$	相伴
$\equiv$	同余
$\lvert \boldsymbol{A} \rvert$	矩阵 $\boldsymbol{A}$ 的行列式, 也记作 $\det \boldsymbol{A}$
$\lvert a \rvert$	绝对值
(a_{ij})	矩阵
$a \mid b$	a 整除 b
(a, b)	最大公因子 (g.c.d.)
$[a, b]$	最小公倍数 (1.c.m.)

索　引

A

阿达玛行列式定理, 286
阿基米德性质, 86
阿基米德定律, 86
埃尔米特矩阵, 262
埃尔米特型, 261
鞍点, 252

B

伴随矩阵, 280
半分配律, 321
半序, 318
半序的包含, 319
半正定, 252
包含, 307
倍立方, 370
被加数, 10
本原多项式, 73
毕达哥拉斯二难推论, 82
变换, 210
变换的乘积, 111
变换的取值域, 210
变换群, 111, 114
标量 (或纯量), 147
标准投影, 180
标准型, 162, 240, 243
标准酉基, 261
标准正交基, 177
并, 312
并立未定元, 62
并-不可约, 324
补子空间, 170
不变量, 240
不变子空间, 297
不变子群, 140
不等式, 7
不可分多项式, 395
不可约多项式, 65, 67
不可约元素, 66
布尔代数, 310
布尔代数的同态, 324
布尔函数多项式, 315
部分分式, 78
悖论, 329
标准型, 217

C

差, 7
长方矩阵, 201
超平面, 166
超越扩张, 360
超越数, 360
抽象代数, 30
抽象群, 117
初等对称多项式, 135
初等矩阵, 212, 277
初等列运算, 216
初等行运算, 157
初等因子, 304
除法算式, 15, 63, 375
传递的关系, 121
传递律, 3, 21, 28
次数, 56

D

戴德金分割, 91
戴德金分割公理, 91
代数闭, 372
代数簇, 350
代数函数域, 364
代数数, 371
代数数域, 373
代数完全, 372
代数整数, 377
代数整数的唯一因子全分解, 383
单扩张, 360
单射, 27
单位, 374, 382
单位根, 98
单位矩阵, 161, 196
单位向量, 153
单位元素, 1, 117, 339
单线性代数, 353
单项矩阵, 199
导数公式, 395
等价, 215, 240
等价关系, 28, 144
等价关系的自反律, 28, 144
等距, 116
笛卡儿, 27
定义关系, 125
定义域, 27
对称差, 322
对称多项式, 398
对称群, 408
对角线法, 371
对角优势矩阵, 279
对偶同构, 402
对偶原理, 319
多项式次数, 53
多项式的唯一因子分解, 70
多项式函数, 55
多项式函数环, 352
多项式理想, 350
多项式形式, 53
多一对应, 27
多重扩张, 368
“度量” 性质, 175

E

二次方程, 5
二次函数的不变式, 249
二次曲线, 245
二次型, 245
二次型标准型, 245
二次型的符号差, 251
二次型的秩, 249
二面体群, 125
二难推论, 82
二项公式, 12
二元关系, 28

F

反对称律, 307
反射, 109, 188, 237
反自同构称, 202
范得蒙, 279
范数, 223, 382
泛界, 309
泛性质, 220
方程, 4
方程的根, 68, 87
仿射变换, 233
仿射几何, 263
仿射空间, 263
仿射群, 234
仿射无关性, 268
仿射子空间, 265
非交换环, 338
非空集合, 9

非空子集, 9
非零因子, 5
非零元素, 5
非奇异的线性变换, 205
非奇异矩阵, 200, 213, 214
非齐次坐标, 271
分割, 90
分划, 144
分块相乘, 203
分配格, 312
分配律, 1, 118, 150
分圆多项式, 76
分支数, 100
辐角, 97
复合数, 19
复平面, 96
复数, 34, 94

G

刚体运动, 239
高斯消去法, 157
高斯引理, 73
高斯整环, 73
高斯整数, 374
高斯整数的唯一因子分解, 375
高斯 (Gauss) 消去法, 40
格, 312, 320
格的同构, 324
格的同态, 324
格拉姆-施密特, 177
根式, 106
根式解法, 106
根域, 385
根域同构, 387
根重数, 101
共轭四元数, 223
共轭直径, 266
共轭子群, 144
构造性证明, 79
关系, 7, 28, 144
归纳假设, 11
G 的同构, 138

H

函数, 26
和, 1, 150
合成, 111
恒等, 56
恒等变换, 112
恒等函数, 56
后继函数, 48
互素, 68, 70
环, 338
环的同态, 60, 341
换位子, 144

J

基数, 327
基数不等式, 331
基数积, 334
基数消去律, 335
积, 1, 27
极大理想, 346
极大值, 252
极小多项式, 291, 317
极小值, 251
集的环, 324
集合, 26, 307
加法群, 233, 380
简化梯形矩阵, 160
减法, 4
交, 312
交错群, 135
交换环, 1, 5, 24, 29, 59, 343
交换律, 1
交换群阿贝耳群, 117
交-不可约, 324

结合代数, 339
解空间, 166
矩阵, 157, 186, 190
矩阵标准型, 162
矩阵乘法, 193
矩阵乘积, 195
矩阵的不变量, 288
矩阵的多项式, 294
矩阵的分解, 299
矩阵的直和, 300
矩阵的秩, 161, 210
矩阵的转置, 202
矩阵的子式, 276
距离, 175
绝对值, 97
伽里略, 329
伽罗瓦群, 392
伽罗瓦域, 389
结合律, 1

K

凯莱定理, 122
凯莱–哈密顿定理, 294, 340
康托对角线法, 331
可除代数, 339, 354
可除环, 343
可分多项式, 395
可分扩张, 395
可解群, 407
可数集, 329
可数集合, 371
克莱姆法则, 281
克罗内克尔符号, 238
克罗内克尔积, 222
空集, 26
扩张, 37

L

拉格朗日定理, 128
拉格朗日 (Lagrange) 插值公式, 58
理想, 342
理想的分解, 383
理想的根式, 353
理想的和, 348
理想的基, 351
理想的积, 349
理想的交, 348
理想论基本定理, 383
理想商, 350
隶莫弗 (De Moivre) 公式, 97
联立同余式, 22
联立线方程组, 215
联立线性方程组, 39, 158
连续统的基数, 331
良序原则, 9
列, 216
列等价, 216
列向量, 202
列运算, 216
列秩, 211
临界点, 252
零, 2
零度, 211
零化子, 183
零矩阵, 192
零空间, 211
零向量, 150
刘维尔数, 374
轮换矩阵, 294
罗伦兹变换, 256
满射, 27
满同态像, 342

M

么模群, 285
幂等律一元运算, 311
幂等元素, 6

幂零矩阵, 292
模, 20
模格, 322
模 n 整数, 25

N

内积, 172
内自同构, 139
逆, 31, 39
逆律, 113, 117

O

欧几里得群, 116, 239
欧几里得算法, 14, 70
欧几里得向量空间, 174
偶数, 6
偶置换, 135

P

判别式, 103
陪集, 128, 129, 180
陪集的积, 142
平行, 265
平行六面体, 283
平行六面体的底, 283
平行四边形法则, 147
平行子空间, 180
平移, 233
谱, 230

Q

奇置换, 135
齐次二次型, 246
齐次线性方程, 42, 165
齐次坐标, 270
切变换, 188, 215
取值域, 27
全线性群, 233
全阵代数, 346
群代数, 339
群的乘法表, 119
群的幂, 123
群的生成元, 124
群的同态, 137
群元素的阶, 124
群 G 的自同构, 138

R

若当标准型, 306
若当矩阵, 305

S

三次方程的三角解法, 89
三次判别式, 105
三等分角, 370
三分律, 8
三角形不等式, 174
三角形矩阵, 199, 278
三重和, 4
商, 14, 31
商环, 345
商空间, 180
商群, 141
商域, 36
上界, 10, 85, 320
射影变换, 272
射影二次曲线, 273
射影几何, 270
射影平面, 270
射影群, 272
射影直线, 270
射影子空间, 271
生成 (或张成), 151
剩余类, 25, 145, 345
施罗德–伯恩斯坦定理, 332
施瓦兹不等式, 174
实对称矩阵, 251, 255, 289
实二次型, 250, 289
实数公设, 85

首项, 56
首项系数, 56
首一多项式, 56
数, 7
数乘运算, 148
数量矩阵, 200
数学归纳法, 10
数学归纳法第二原理, 12
数学归纳法原理, 10
数学系统, 1
"数乘" 积, 147, 193
双边理想, 353
双边一般分配律, 11
双射, 27
双线性, 172
双线性标准型, 245
双线性函数, 218
双线性型, 244
双性性标准型的秩, 245
素, 14
素理想, 346
素数, 14
素因子, 14
算术基本定理, 18

T

特征, 356
特征多项式, 288
特征方程, 288
特征根, 230, 288
特征向量, 287
特征值, 230, 288
梯形矩阵, 160
体, 223
同构, 29, 56
同态, 60
同态的核, 137
同一律, 112, 118
同余, 20
同余关系, 21, 144
同余式, 20
T- 循环子空间, 297

W

完备的有序域, 85
完全可约, 299
唯一性, 1
唯一因子分解, 69
唯一因子分解整环, 72
稳定型方程, 107
无理数, 82
无限集, 329
无限小数, 84
五次不可解性, 406

X

析取标准型, 316
吸收律, 311
希尔伯特基定理, 352
希尔伯特零点, 353
下界, 10, 85, 320
线性变换, 188, 210, 299
线性代数的阶, 338
线性的, 39
线性方程, 39
线性函数, 181
线性内插法, 89
线性式变换, 272
线性无关, 153
线性相关, 154
线性型, 242, 243
相伴, 65
相等关系, 2
相等关系的定律, 21
相容律原理, 309
相似变换, 239
相似矩阵, 230

向量长, 261
向量的坐标, 229
消去律, 3, 5
消元法, 41
斜对称, 260
斜对称矩阵, 246
斜线性, 261
形式导数, 395
形心, 266
行, 157
行等价, 162, 213
行简化, 160
行简化矩阵, 160
行矩阵, 202
行空间, 157
行列式, 40, 275, 279
行列式的积, 279
行列式秩, 281
行秩, 211
性组合, 16
虚分量, 94
序-同构, 50
旋转, 109, 186
玄, 4
选择公理, 113, 329
循环群, 124
循环置换循环, 131

Y

一般分配律, 10
一般交换律, 11
一般结合律, 11
一一变换, 112
一一对应, 27
酉变换, 261
酉基, 261
酉矩阵, 261
酉空间, 261
有理标准型, 304
有理函数, 57, 352
有理式, 54
有理数, 35
有理整数, 377
有限集, 329
有限扩张, 365
有限维的, 155
有序域, 43, 83
有序整环, 7, 85
友矩阵, 292
右分配律, 3
右理想, 353
右逆元素, 112, 118
右 (左) 陪集指数, 129
余数, 15, 64
余数定理, 64
余子式, 276
域, 32, 118
域扩张的生成元, 360
元数学, 319
原子, 324
运算, 27
运算微积, 196

Z

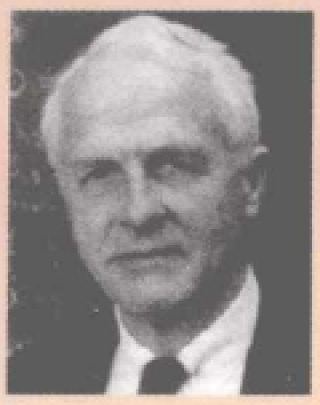
增广矩阵, 215
张量积, 218, 221
真包含, 308
真理想, 343
真满同态, 343
真子群, 126
整环, 1, 2, 5
整数, 1
正, 7
正定, 252
正方形的对称, 109
正规子群, 140, 402

正交变换, 237
正交矩阵, 238
正交群, 238
正交向量, 261
正元素, 7
正整数, 1, 7
正整数公设, 45
正 (交) 投影, 178
直和, 169
指数, 11
置换, 114, 131
置换矩阵, 199
置换群, 131
中点, 264
重心坐标, 268
主理想, 69
主轴, 254
主轴定理, 254
准素分支, 300
准素有理标准型, 304
子代数, 316
子代数的生成, 315
子格, 324
子环, 60, 338
子集, 26, 307
子矩阵, 203, 276
子空间, 151, 152
子空间的交, 152
子空间的像, 209
子空间的直和, 169
子群, 126, 129
子群的交, 127
子域, 33
子整环, 5
自反律, 2, 28, 307
自共轭子群, 140
自然倍数, 355
自同构, 29, 205
自由布尔代数, 323
最大公因子, 70, 320
最大下界, 83, 320
最大元素, 10
最小公倍数, 16, 348
最小上界, 83, 320
最小元素, 10
左单位元素, 118
左分配律, 3
左理想, 353
左逆元素, 112
坐标, 168
转置, 190

近世代数概论（第5版）
A Survey of Modern Algebra

近世代数也称抽象代数，是现代数学的重要基础，主要研究群、环、域等代数结构。它的概念与思想渗透到所有数学分支，而其理论与方法在统计学、信息论、计算机科学、近代物理、化学以及其他许多科学与工程领域中都有广泛而深入的应用。

这本经典的教材出自抽象代数领域的两位巨匠之手，曾对近世代数教学产生深远的影响，帮助了几代学子理解和掌握近世代数，至今本书仍是一部对自学和课堂教学都极具价值的参考书和教材。作者用大家熟悉且具体的例子来阐述每一个概念，深入浅出，透彻简洁。为了培养学生独立思考的能力，每个专题都包括丰富的练习。



Garrett Birkhoff（1911—1996）已故世界著名数学家。生前曾任国际数学家大会组织委员会主席、美国数学会副主席、美国工业与应用数学会主席、《大不列颠百科全书》编委，美国科学院院士，哈佛大学教授。1933年开创格论研究，使其成为数学的一个重要分支。



Saunders Mac Lane（1909—2005）已故世界著名数学家。生前曾任美国数学会(MAA)副主席、主席，美国数学协会(AMS)副主席、主席，美国科学院副院长、院士。Mac Lane的贡献主要在代数和代数拓扑方面，他是同调代数和范畴论的先驱者之一，因在代数及代数拓扑方面的贡献获1986年美国数学会斯蒂尔奖。1989年获美国科学界最高荣誉国家科学奖。

本书相关信息请访问：
图灵网站 http://www.turingbook.com
读者/作者热线：(010) 88593802
反馈/投稿/推荐信箱： contact@turingbook.com

分类建议 数学／基础数学

人民邮电出版社网址 www.ptpress.com.cn

ISBN 978-7-115-18387-3

9 787115 183873 >

ISBN 978-7-115-18387-3/O1

定价：69.00元